

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:47:31; Search time 103.249 Seconds

(without alignments)
371.762 Million cell updates/sec

Title: US-10-054-873-6

Perfect score: 587

Sequence: 1 MPTPLSLRFLFDNMLRAHR.....IVEQCCTSLICSLYLENYCN 107

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_23Sep04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003ds:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	587	100.0	107	2	AAV42860 hGH-mini-
2	555.5	94.6	150	2	AAV42861 Chimeric
3	315.5	53.7	116	2	AAV42862 SOD-protein
4	304	51.8	63	2	AAV42863 Human pro
5	304	51.8	117	2	AAV42864 SOD-protein
6	302.5	51.5	137	2	AAV42865 SOD-protein
7	299	50.9	56	2	AAV42866 Matrig fa
8	299	50.9	56	2	AAV42867 Human pro
9	299	50.9	96	2	AAV42868 Proinsulin
10	299	50.9	96	2	AAV42869 Human pro
11	299	50.9	145	2	AAV42870 Fusion pr
12	299	50.9	145	2	AAV42871 Matrig fa
13	299	50.9	145	2	AAV42872 Matrig fa
14	293	49.9	52	2	AAV42873 Human ins
15	288.5	49.1	160	2	AAV42874 Proinsulin
16	287	48.9	52	2	AAV42875 Glycosylp
17	287	48.9	52	2	AAV42876 Example o
18	287	48.9	52	2	AAV42877 Peptidins
19	286	48.7	138	2	AAV42878 PIV142 mo
20	286	48.7	138	2	AAV42879 Human alb
21	285	48.6	116	8	AAV42880 Human alb
22	284.5	48.5	58	2	AAV42881 Prepro-in
23	284.5	48.5	59	2	AAV42882 Modified
24	284.5	48.5	65	2	AAV42883 N-termina
25	284.5	48.5	109	1	AAV42884 Amino aci

26	284.5	48.5	123	2	AAV42885
27	284.5	48.5	123	2	AAV42886
28	284.5	48.5	124	2	AAV42887
29	284.5	48.5	124	6	AAV42888
30	284.5	48.5	124	6	AAV42889
31	284.5	48.5	125	2	AAV42890
32	284	48.4	116	8	AAV42891
33	284	48.4	118	8	AAV42892
34	284	48.4	118	2	AAV42893
35	284	48.4	118	2	AAV42894
36	284	48.4	140	2	AAV42895
37	284	48.4	671	7	AAV42896
38	284	48.4	671	7	AAV42897
39	284	48.4	671	7	AAV42898
40	284	48.4	671	7	AAV42899
41	283.5	48.3	53	2	AAV42900
42	283.5	48.3	53	2	AAV42901
43	283.5	48.3	117	2	AAV42902
44	283.5	48.3	408	4	AAV42903
45	283	48.2	667	7	AAV42904

ALIGNMENTS

RESULT 1	AAV42860	standard; protein; 107 AA.
ID	AAV42860	
XX	AAV42860;	
AC	19-JAN-2000	(first entry)
DT	19-JAN-2000	
XX	hGH-mini-proinsulin chimeric protein.	
DE	Insulin; precursor; growth hormone; chaperone; intramolecular; folding; conformation; chimeric protein; cleavable; recombinant; production; yield.	
KW	Synthetic.	
OS	Homo sapiens.	
XX	WO950302-A1.	
XX	07-OCT-1999.	
PD	31-MAR-1998;	98WO-CN000052.
PF	31-MAR-1998;	98WO-CN000052.
XX	(TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.	
PA	Gan Z;	
PI	WPI; 1999-610839/52.	
DR	New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.	
PT	Claim 13; Page 30; 46pp; English.	
PS	This sequence represents a chimeric protein, hGH-mini-proinsulin. This chimeric protein contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AAV42855, a cleavable peptide linker (AAV42857), and a human insulin precursor comprising insulin A and B chains (AAV42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only	

CC protect insulin sequences from intracellular degradation by a
 CC microorganism host, but also promote the folding of the fused insulin
 CC precursor, facilitate the solubility of the fusion protein and decrease
 CC the intermolecular interactions among the fusion proteins, thus allowing
 CC folding of the fused insulin precursor at commercially useful high
 CC concentrations. The procedural steps of cyanogen bromide cleavage,
 CC oxidative sulphytolysis and related purification steps can thus be
 CC eliminated, along with the use of high concentrations of mercapran or the
 CC use of hydrophobic absorbent resins

XX Sequence 107 AA:

Query Match 100.0%; Score 587; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3e-43;
 Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSRPLDNNMLRAHRLHQLAFTYQEFEEAVYIPKQKXSFQNPGLTGPRFVNOH 60
 Db 1 MFPTPLSRPLDNNMLRAHRLHQLAFTYQEFEEAVYIPKQKXSFQNPGLTGPRFVNOH 60
 QY 61 LGGSHLVEALYLVCGERGFPYTPKTRGIVEOCCSTICSLYQLENVCN 107
 Db 61 LGGSHLVEALYLVCGERGFPYTPKTRGIVEOCCSTICSLYQLENVCN 107

RESULT 2

ID AAY42861 standard; protein; 150 AA.

XX AAY42861;

XX 19-JAN-2000 (first entry)

XX Chimeric protein, SEQ ID 7.

XX Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
 XX conformation; chimeric protein; cleavable; recombinant; production;
 XX yield.

XX Synthetic.

XX Homo sapiens.

XX WO9950302-A1.

XX 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN000052.

XX 31-MAR-1998; 98WO-CN000052.

XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX Gan Z;

XX WPI; 1999-610839/52.

XX New chimeric proteins containing human growth hormone fragment, used
 XX particularly for the production of human insulin.

XX Claim 14; Page 30-31; 46pp; English.

XX This sequence represents a chimeric protein, which contains an N-terminal
 CC fragment of human growth hormone (hGH) of the sequence given in AAY42856,
 CC a cleavable peptide linker (AAY42857), and a human insulin precursor
 CC comprising insulin A and B chains (AAY42859). The hGH portion of the
 CC chimeric protein acts as an intramolecular chaperone (IMC) for the
 CC insulin precursor, enabling it to fold correctly. The cleavable peptide
 CC linker has a C-terminal Arg residue which enables the hGH portion of the
 CC chimeric protein to be removed after folding has taken place. Production
 CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
 CC provide human insulin with correctly linked cysteine bridges with fewer
 CC necessary procedural steps, and hence resulting in a higher yield of
 CC human insulin. The IMC sequences not only protect insulin sequences from

CC intracellular degradation by a microorganism host, but also promote the
 CC folding of the fused insulin precursor, facilitate the solubility of the
 CC fusion protein and decrease the intermolecular interactions among the
 CC fusion proteins, thus allowing folding of the fused insulin precursor at
 CC commercially useful high concentrations. The procedural steps of cyanogen
 CC bromide cleavage, oxidative sulphytolysis and related purification steps
 CC can thus be eliminated, along with the use of high concentrations of
 CC mercapran or the use of hydrophobic absorbent resins

XX Sequence 150 AA:

Query Match 94.6%; Score 555.5; DB 2; Length 150;
 Best Local Similarity 71.3%; Pred. No. 2.1e-40;
 Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MFPTPLSRPLDNNMLRAHRLHQLAFTYQEFEEAVYIPKQKXSFQNPGLTGPRFVNOH 49
 Db 1 MFPTPLSRPLDNNMLRAHRLHQLAFTYQEFEEAVYIPKQKXSFQNPGLTGPRFVNOH 60
 QY 50 -----LGTGPRFVNOHLCGSHLVEALYLVCGER 77
 Db 61 TFSNEETQKSNELLRLISILLIQSWLEPVQGTGPRFVNOHLCGSHLVEALYLVCGER 120
 QY 78 GFFYTPKTRGIVEOCCSTICSLYQLENVCN 107
 Db 121 GFFYTPKTRGIVEOCCSTICSLYQLENVCN 150

RESULT 3

ID AAR98897 standard; protein; 116 AA.

XX AAR98897;

XX 03-FEB-1997 (first entry)

XX SOD-proinsulin hybrid polypeptide.

XX Insulin; proinsulin; hybrid polypeptide; protein folding;

XX enzymatic cleavage; cyanogen bromide; sulphytolysis.

XX Homo sapiens.

XX WO9620724-A1.

XX 11-JUL-1996.

XX 29-DEC-1994; 94WO-US013268.

XX 29-DEC-1994; 94WO-US013268.

XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.

XX Hartman JR, Mendelovitz S, Gorecki M;

XX WPI; 1996-333766/33.

XX N-PSDB; AAT34670.

XX Recombinant insulin prodn. by correctly folding pro-insulin hybrid
 XX polypeptide - then enzymatic cleavage of folded product, does not require
 XX sulphite protection of SH nor use of cyanogen bromide.

XX Example 1B; Fig 7; 69pp; English.

XX A new method for the production of recombinant human insulin comprises
 CC folding a hybrid polypeptide comprising proinsulin under conditions that
 CC permit correct disulphide bond formation and subjecting that folded
 CC protein to enzymatic cleavage. The insulin produced can then be purified.
 CC This sequence is a SOD-insulin B chain-Arg-insulin A chain hybrid.
 CC Polypeptide and is encoded by the plasmid construct pDBAST-LAT.
 CC Transformation of the proper E.coli host cells with pDBAST-LAT results in
 CC the efficient expression of the proinsulin hybrid polypeptide, useful for
 CC human insulin production. The method produces recombinant human insulin

RESULT 6

```

AA07692 standard; protein; 137 AA.
ID   AAR71692
XX AC   AAR71692;
XX DT   25-MAR-2003 (revised)
XX DT   20-NOV-1995 (first entry)
XX DE   Mating factor alpha 1-Insulin precursor Arg831.
XX KM   Human insulin precursor Arg831; diabetes; zinc ion complex;
XX KW   mating factor alpha 1.
XX OS   Homo sapiens.
XX PH   Key
FH Protein      Location/Qualifiers
FT             1..85
FT             /label= mating factor alpha-1
FT Peptide     86..116
FT             /label= B-chain
FT Peptide    117..137
FT             /label= A-chain
XX PN   WO507931-A1.
XX PD   23-MAR-1995.
XX PF   16-SEP-1994; 94NO-DK000347.
XX PR   17-SEP-1993; 93DK-00001044.
XX PR   02-FEB-1994; 94US-00190829.
XX PA   (NOVO ) NOVO-NORDISK AS.
PI Havelund S, Halstrom JB, Jorassen I, Andersen AS, Markusen U;
DR WP1: 1995-131314/17.
DR N-PSDB: AA086425.
XX PT   Acylated insulin deriv. which may be present as a Zinc ion complex - is
XX FT   used to treat diabetes and is rapid acting.
XX PS   Example 5; Page 78; 100pp; English.
XX CC   AA086425 encodes AAR71692 mating factor alpha 1-Insulin precursor Arg831.
XX CC   AARB31 comprises the B and A chains of a claimed human insulin
XX CC   derivative. In the final claimed compsn they are covalently connected
XX CC   via disulphide bonds between Cys residues A1/B7 and A20/B15. The
XX CC   derivative, which may be present as a zinc ion complex, can be used as a
XX CC   fast action treatment for diabetes. (Updated on 25-MAR-2003 to correct PN
XX CC   field.)
XX SC   Sequence 137 AA:
SQ
Query Match          51.5%; Score 302.5; DB 2; Length 137;
Best Local Similarity 50.0%; Pred. No.1.3e-18;
Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4
QY 2 PPTPLBSRLFDNMLRAHRLQLADTYQEPPEAVIPKEQ--KISFLQ-----N 48
DQ 3 FPST-----FTAVLFAASALAPVNVTTEDETAQIPEAVIGSDLEGDFDAVLHFSN 57
QY 49 FLNG-----PFFVQHLCGSHLYEATLVNGEGFFYTPTKG 87
DB 58 STNNGLLFINTTIAIAKEBGVMNAKFVNQHLCGSHLYEATLVNGEGGFYTPTXRG 117
QY 88 IVBOCCSTICSLYLENYCN 107
DB 118 IVBOCCSTICSLYLENYCN 137

```

ID	AA68930	standard; peptide; 56 AA.
AC	AA68901:	
XX		
DT	25-MAR-2003	(revised)
DT	02-MAR-1995	(first entry)
XX		
DE	Human pro-insulin 3.	
XX		
KM	Pro-insulin; A-chain; B-chain; C-chain; disulphide; mercaptan;	
KM	chaotropic agent.	
XX		
OS	Homo sapiens.	
XX		
PN	EP600372-A1.	
XX		
XX	08-JUN-1994.	
XX		
XX	25-NOV-1993;	93EP-00118993.
XX		
XX	02-DEC-1992;	92DE-04240420.
XX		
PA	(FARH) HOECHST AG.	
XX		
PI	Obermeier R, Gerl M, Ludwig J, Sabel W;	
XX		
DR	WPI; 1994-177718/22.	
XX		
PT	Prodn. of pro-insulin with correct di:sulphide bridges - by treating	
PT	recombinant precursor protein with mercaptan in alkali and in presence of	
PT	chaotropic agent, then isolation on hydrophobic resin.	
XX		
PS	Disclosure; Page 12; 15pp; German.	
XX		
CC	Pro-insulin is produced by treating recombinant precursor protein with a	
CC	mercaptan to provide 2-10 SH residues per Cys residue, in presence of a	
CC	chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3	
CC	-50 g hydrophobic adsorber resin per 1 aq. medium of pH 4-7, isolating	
CC	the adsorbed resin and pro-insulin and desorbing the pro-insulin. This	
CC	method produces pro-insulin with correctly bonded Cys bridges. Compared	
CC	with known methods it involves fewer stages (esp. no sulphytolysis or	
CC	cyanogen bromide cleavage) and overall losses during purification are	
CC	reduced, i.e. the process is quicker and gives better yields. Sequences	
CC	of insulin chain A, B and C are given in AA68895-97. Sequences of pro-	
CC	insulin 1-4 are given in AA68898-901. (Updated on 25-MAR-2003 to correct	
CC	PN field.)	
XX		
SO	Sequence 56 AA:	
XX		
Query Match	50.9%;	Score 299; DB 2; Length 56;
Best Local Similarity	100.0%;	Pred. No. 1,2e-18;
Matches 53; Conservative	0;	Mismatches 0; Indels 0; Gaps 0
55	RFVNGQLCGSHLYEALVYVCGRGFFYTPKRGIVGECCTSIQVLYENYCN	107
4	RFVNGQLCGSHLYEALVYVCGRGFFYTPKRGIVGECCTSIQVLYENYCN	56
RESULT 8		
AA68895		
ID	AA68895 standard; protein; 56 AA.	
XX		
AC	AA68865;	
XX		
DT	03-APR-1996	(first entry)
XX		
DE	Proinsulin sequence 3.	
XX		
KM	Proinsulin; post-translational modification; recombinant production;	
KM	protein folding; conformation.	
XX		
OS	Synthetic.	
XX		

FH Key Location/Qualifiers
 FT Region 1..4
 FT /label= R2
 FT /note= "a peptide of 4 amino acids"
 FT Peptide 5..34
 FT /label= R1-(B2-B29)-Y
 FT /note= "human insulin B-chain"
 FT Region 35
 FT /label= X
 FT 36..56
 FT /label= Gly-(A2-A20)-R3
 FT /note= "human insulin A-chain"
 XX
 XX EP66892-A2.
 XX
 XX 23-AUG-1995.
 XX
 XX 09-FEB-1995; 95EP-00101748.
 XX
 XX 18-FEB-1994; 94DE-04405179.
 XX
 XX (FARR) HOECHST AG.
 XX
 XX Obermeier R, Gerl M, Ludwig J, Sabel W;
 XX
 XX WPI; 1995-284754/38.
 XX
 XX Isolation of insulin that is correctly post-translationally processed -
 PT by reacting proinsulin with a mercaptan in the presence of a chaotropic
 PT agent and purification after absorption to hydrophobic resin.
 XX
 XX Example 2; Page 13; 16pp; German.
 XX
 XX The present sequence is an example of a proinsulin molecule corresp. to
 CC the general formula R2-R1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (II). In formula
 CC (II), X = Lys, Arg or a peptide of 2-35 amino acids contg. Lys or Arg at
 CC the N- and C-termini; Y = a natural amino acid; R1 = Phe or a bond; R2 =
 CC H, Arg, Lys, a peptide of 2-45 amino acids contg. Arg or Lys at the N-
 CC and C-termini; R3 = a natural amino acid; (A2-A20) and (B2-B29) are the
 CC insulin A- and B-chain sequences from human or other insulin. The
 CC proinsulin molecule (produced in recombinant E. coli) is reacted with
 CC mercaptan at a ratio of 2-10 SH residues of mercaptan per Cys residue of
 CC proinsulin. The reaction takes place in the presence of a chaotropic
 CC auxiliary agent at pH 10-11 and results in proinsulin with correctly
 CC linked cystine bridges. Reaction with trypsin and opt. carboxypeptidase B
 CC yields correctly folded insulin. The insulin is isolated by absorption on
 CC a hydrophobic resin
 XX
 XX Sequence 56 AA;
 SO
 Query Match 50.9%; Score 299; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 55 RPNVNHCGSHLVEALYLVCGERGFFYPTKTRIGIVECCCTISCSLYQLENYCN 107
 DB 4 RPNVNHCGSHLVEALYLVCGERGFFYPTKTRIGIVECCCTISCSLYQLENYCN 56
 RESULT 9
 AAR68899 standard; peptide; 96 AA.
 ID AAR68899;
 XX
 XX AAR68899;
 XX
 XX 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 XX Human pro-insulin 2.
 XX
 XX Pro-insulin; A-chain; B-chain; C-chain; disulphide; mercaptan;
 KM chaotropic agent.
 XX

OS Homo sapiens.
 XX
 XX EP600372-A1.
 XX
 XX 08-JUN-1994.
 PD
 XX
 XX 25-NOV-1993; 93EP-00118993.
 XX
 XX 02-DEC-1992; 92DE-04240420.
 XX
 XX (FARR) HOECHST AG.
 XX
 XX Obermeier R, Gerl M, Ludwig J, Sabel W;
 XX
 XX WPI; 1994-177718/22.
 DR
 XX
 XX Prodn. of pro-insulin with correct disulphide bridges - by treating
 PT recombinant precursor protein with mercaptan in alkali and in presence of
 PT chaotropic agent, then isolation on hydrophobic resin.
 XX
 XX Disclosure; Page 11; 15pp; German.
 PS
 XX
 XX Pro-insulin is produced by treating recombinant precursor protein with a
 CC mercaptan to provide 2-10 SH residues per Cys residue, in presence of a
 CC chaotropic agent and in aq. medium of pH 10-11, treating the prod. with 3
 CC -50 g hydrophobic adsorbent resin per 1 aq. medium of pH 4-7, isolating
 CC the adsorbed resin and pro-insulin and desorbing the pro-insulin. This
 CC method produces pro-insulin with correctly bonded Cys bridges. Compared
 CC with known methods it involves fewer stages (esp. no sulphiteolysis or
 CC cyanogen bromide cleavage) and overall losses during purification are
 CC reduced, i.e. the process is quicker and gives better yields. Sequences
 CC of insulin chain A, B and C are given in AAR68895-97. Sequences of pro-
 CC insulin 1-4 are given in AAR68898-901. (updated on 23-MAR-2003 to correct
 CC PN field.)
 CC
 XX
 XX Sequence 96 AA;
 SO
 Query Match 50.9%; Score 299; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 2e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 55 RPNVNHCGSHLVEALYLVCGERGFFYPTKTRIGIVECCCTISCSLYQLENYCN 107
 DB 44 RPNVNHCGSHLVEALYLVCGERGFFYPTKTRIGIVECCCTISCSLYQLENYCN 96
 RESULT 10
 AAR78662 standard; protein; 96 AA.
 ID AAR78662;
 XX
 XX AAR78662;
 XX
 XX 03-APR-1996 (first entry)
 DT
 XX
 XX Fusion protein contg. proinsulin sequence 3.
 DE Proinsulin; post-translational modification; recombinant production;
 KM protein folding; conformation.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Region 41..44
 FT /label= R2
 FT /note= "a peptide of 4 amino acids"
 FT Peptide 45..74
 FT /label= R1-(B2-B29)-Y
 FT /note= "human insulin B-chain"
 FT Region 75
 FT /label= X
 FT 76..96
 FT /label= Gly-(A2-A20)-R3
 FT /note= "human insulin A-chain"
 FT

XX PP668292-A2.
 PN 23-AUG-1995.
 XX 09-FEB-1995; 95EP-00101748.
 XX 18-FEB-1994; 94DE-04405179.
 XX (FARH) HOECHST AG.
 XX Obermeier R, Gerl M, Ludwig J, Sabel W;
 DR WPI; 1995-284754/38.
 XX
 PT Isolation of insulin that is correctly post-translationally processed
 PT by reacting pro-insulin with a mercaptan in the presence of a chaotropic
 PT agent and purification. after absorption to hydrophobic resin.
 XX
 PS Example 2; Page 8; 16pp; German.
 XX
 CC The present sequence is that of a fusion protein, produced in E.coli
 CC which contains an example of a proinsulin molecule corresp. to the
 CC general formula R2-R1-(B2-B29)-Y-X-Gly-(A2-A20)-R3 (II). In formula (II),
 CC X = Lys, Arg or a peptide of 2-35 amino acids contg. Lys or Arg at the N-
 CC and C-terminal; Y = a natural amino acid; R1 = Phe or a bond; R2 = H, Arg,
 CC Lys, a peptide of 2-45 amino acids contg. Arg or Lys at the N- and C-
 CC terminal; R3 = a natural amino acid; (A2-A20) and (B2-B29) are the insulin
 CC A- and B-chain sequences from human or other insulin. The proinsulin
 CC molecule, released by cyanogen bromide, is reacted with mercaptan at a
 CC ratio of 2:10 SH residues of mercaptan per Cys residue of proinsulin. The
 CC reaction takes place in the presence of a chaotropic auxiliary agent at
 CC pH 10-11 and results in proinsulin with correctly linked cysteine bridges.
 CC Reaction with trypsin and opt. carboxypeptidase B yields correctly folded
 CC insulin. The insulin is isolated by absorption on a hydrophobic resin
 CC
 SQ Sequence 96 AA;
 Query Match 50.9%; Score 299; DB 2; Length 96;
 Best Local Similarity 100.0%; Pred. No. 28-18; Indels 0; Gaps 0;
 Matches 53; Conservative 0; Mismatches 0;
 QY 55 RFVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVGQCTSIICSLYLENNYCN 107
 Db 44 RFVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVGQCTSIICSLYLENNYCN 96
 RESULT 11
 ID AAR71694 standard; protein: 145 AA.
 AC AAR71694;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-NOV-1995 (first entry)
 XX
 DE Mating factor alpha 1-insulin precursor ArgB1, ArgB31 N-terminal.
 XX
 KM Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
 KM mating factor alpha 1; N-terminal EBAEAEAR.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Protein
 FT 1..85 Location/Qualifiers
 FT /label= mating factor alpha-1
 FT Peptide
 FT 86..93
 FT /label= N-terminal peptide
 FT Peptide
 FT 94..124
 FT /label= B-chain
 FT 125..145
 FT /label= A-chain
 FT Peptide
 XX

PN WO9507931-A1.
 XX
 PD 23-MAR-1995.
 XX
 PF 16-SEP-1994; 94WO-DK000347.
 XX
 PR 17-SEP-1993; 93DK-00001044.
 PR 02-FEB-1994; 94US-00190829.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 PI Havelund S, Halstrom JB, Jonassen I, Andersen AS, Markussen J;
 DR WPI; 1995-133114/17.
 DR N-PSDB; AA086429.
 XX
 PT Acylated insulin deriv. which may be present as a zinc ion complex - is
 PT used to treat diabetes and is rapid acting.
 XX
 PS Example 5; Page 82-83; 100pp; English.
 XX
 CC AA086429 encodes AAR71694 mating factor alpha 1-insulin precursor ArgB1.
 CC ArgB31 N-terminal EBAEAEAR. The insulin precursor comprises the B and A
 CC chains of a claimed human insulin derivative preceded by the N-terminal
 CC amino acids EBAEAEAR. In the final claimed compsn. they are covalently
 CC connected via disulphide bonds between Cys residues A1/B7 and A20/B19.
 CC The derivative, which may be present as a zinc ion complex, can be used
 CC as a fast action treatment for diabetes. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 145 AA;
 Query Match 50.9%; Score 299; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 28-18; Indels 0; Gaps 0;
 Matches 53; Conservative 0; Mismatches 0;
 QY 55 RFVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVGQCTSIICSLYLENNYCN 107
 Db 93 RFVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVGQCTSIICSLYLENNYCN 145
 RESULT 12
 ID AAR71695 standard; protein: 146 AA.
 AC AAR71695;
 XX
 DT 25-MAR-2003 (revised)
 DT 20-NOV-1995 (first entry)
 XX
 DE Mating factor alpha 1-insulin precursor ArgB1, ArgB31 N-terminal.
 XX
 KM Human insulin precursor ArgB1, ArgB31; diabetes; zinc ion complex;
 KM mating factor alpha 1; N-terminal EBAEAEAR.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FH Protein
 FT 1..85 Location/Qualifiers
 FT /label= mating factor alpha-1
 FT Peptide
 FT 86..94
 FT /label= N-terminal peptide
 FT Peptide
 FT 95..125
 FT /label= B-chain
 FT 126..146
 FT /label= A-chain
 PN WO9507931-A1.
 XX
 PD 23-MAR-1995.
 XX
 PF 16-SEP-1994; 94WO-DK000347.
 XX

PR 17-SEP-1993: 93DK-00001044.
 PR 02-FEB-1994: 94US-00190829.
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX
 PI Havelund S, Halstrom JB, Jonassen I, Andersen AS, Markussen J,
 XX
 DR WPI; 1995-131314/17.
 DR N-PSDB; AAQ86432.
 XX
 PT Acylated insulin deriv. which may be present as a zinc ion complex - is
 PT used to treat diabetes and is rapid acting.
 XX
 XX Example 6; Page 85; 100pp; English.
 XX
 CC AAQ86432 encodes AAR71695 mating factor alpha 1-insulin precursor Arg31,
 CC Arg31 N-terminal ERAEAAER. The insulin precursor comprises the B and A
 CC chains of a claimed human insulin derivative preceded by the N-terminal
 CC amino acids ERAEAAER. In the final claimed compn, they are covalently
 CC connected via disulphide bonds between Cys residues A7/B7 and A28/B19.
 CC The derivative, which may be present as a zinc ion complex, can be used
 CC as a fast action treatment for diabetes. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 CC
 SQ Sequence 146 AA;
 XX
 QY Query Match 50.9%; Score 299; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 2,9e-18;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 55 FVNHQCGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIQSLYLENYCN 107
 94 RFVNHQCGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIQSLYLENYCN 146
 XX
 RESULT 13
 AA42859
 ID AAY42859 standard; protein; 52 AA.
 AC
 AC AAY42859;
 DT 19-JAN-2000. (first entry)
 XX
 DE Human insulin precursor, SEQ ID 5.
 XX
 KM Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
 KM conformation; chimeric protein; cleavable; recombinant; production;
 KM yield.
 XX
 OS Homo sapiens.
 XX
 PN WO950302-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 31-MAR-1998; 98WO-CN000052.
 XX
 PR 31-MAR-1998; 98WO-CN000052.
 XX
 XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX
 PI Gan Z;
 XX
 DR WPI; 1999-610839/52.
 XX
 PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.
 XX
 PS Claim 12; Page 29-30; 46pp; English.
 XX
 CC This sequence represents a human insulin precursor comprising insulin A
 CC and B chains. This insulin precursor is a component of the chimeric
 CC proteins hGH-mini-proinsulin (AAY42860) and the chimeric protein given in

CC AAY42861. These chimeric proteins additionally contain an N-terminal
 CC fragment of human growth hormone (hGH) and a cleavable peptide linker
 CC (AAY42857). The hGH portion of the chimeric protein acts as an
 CC intramolecular chaperone (IMC) for the insulin precursor, enabling it to
 CC fold correctly. The cleavable peptide linker has a C-terminal Arg residue
 CC which enables the hGH portion of the chimeric protein to be removed after
 CC folding has taken place. Production of recombinant human insulin via an
 CC hGH-proinsulin chimeric protein can provide human insulin with correctly
 CC linked cysteine bridges with fewer necessary procedural steps, and hence
 CC resulting in a higher yield of human insulin. The IMC sequences not only
 CC protect insulin sequences from intracellular degradation by a
 CC microorganism host, but also promote the folding of the fused insulin
 CC precursor, facilitate the solubility of the fusion protein and decrease
 CC the intermolecular interactions among the fusion proteins, thus allowing
 CC folding of the fused insulin precursor at commercially useful high
 CC concentrations. The procedural steps of cyanogen bromide cleavage,
 CC oxidative sulphydrololysis and related purification steps can thus be
 CC eliminated, along with the use of high concentrations of mercaptan or the
 CC use of hydrophobic absorbent resins
 CC
 SQ Sequence 52 AA;
 XX
 QY Query Match 50.1%; Score 294; DB 2; Length 52;
 Best Local Similarity 100.0%; Pred. No. 3e-18;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 56 FVNHQCGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIQSLYLENYCN 107
 1 FVNHQCGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIQSLYLENYCN 52
 XX
 RESULT 14
 AAR04582
 ID AAR04582 standard; protein; 57 AA.
 AC
 AC AAR04582;
 DT 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 14-SEP-1990 (first entry)
 XX
 DE Proinsulin analogue with a Lys residue linking the A and B chains.
 XX
 KM Insulin fusion protein; pro-insulin analogue; tandemstate;
 KM Lys-Lys bridge; ds.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..35
 FT /note= "Insulin B chain"
 FT /note= "Lys residue linking insulin B chain to A chain"
 FT Peptide 37..57
 FT /note= "Insulin A chain"
 XX
 EN EP367163-A.
 XX
 ED 09-MAY-1990.
 XX
 PF 28-OCT-1989; 89EP-00120056.
 XX
 PR 03-NOV-1988; 88DE-03837273.
 PR 19-AUG-1989; 89DE-03927449.
 XX
 XX (FARR) HOECHST AG.
 XX
 PI Koller KP, Riess GJ, Uhlmann E, Wellmeier H;
 XX
 DR WPI; 1990-141149/19.
 DR N-PSDB; AAC04335.
 XX
 PT New insulin fusion proteins - comprise pro-insulin analogue linked to

PT tendamistate.
 XX
 PS Disclosure; Page 5; 8pp; German.
 XX
 CC This sequence is joined to the C-terminus of an N-terminal fragment
 CC comprising opt. modified tendamistate. This fusion protein may be
 CC converted into human insulin using known methods. The synthetic gene was
 CC prepared by the phosphoramidite method. See also AA004336. (Updated on 25
 CC -MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI
 CC field.)
 CC
 CC Revised record issued on 09-SEP-2004 : Correction to pages and features
 XX
 SQ Sequence 57 AA;
 Query Match 49.8%; Score 293; DB 2; Length 57;
 Best Local Similarity 96.2%; Pred. No. 4e-18; 0; Indels 0; Gaps 0;
 Matches 51; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 55 RFVNOHLGSHLVEALYLVGGERGFFYTPKTRGIVEQCCTSIQSLYLENYCN 107
 5 KFVNOHLGSHLVEALYLVGGERGFFYTPKTRGIVEQCCTSIQSLYLENYCN 57
 DB
 RESULT 15
 AAR79055
 ID AAR79056 standard; protein: 160 AA.
 XX
 AC AAR79056;
 XX
 DT 25-MAR-2003 (revised)
 DT 24-JAN-1996 (first entry)
 XX
 DE Glycosylphosphatidylinositol-anchored human recombinant insulin.
 XX
 KM GPI: glycosylphosphatidylinositol; insulin; hormone; solubilization;
 KM Saccharomyces cerevisiae; anchor; Gasi; plasmid pBY40.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 44..129
 FT /note="anchor attachment site"
 XX
 PN WO9552614-A1.
 XX
 PD 24-AUG-1995.
 XX
 PF 16-FEB-1995; 95WC-BR000010.
 XX
 PR 17-FEB-1994; 94BR-0000600.
 XX
 BA (FINE-) FINEP FINANCIADORA ESTUDOS & PROJETOS.
 BA (ESCO-) ESCOLA PAULISTA MEDICINA.
 PI Cardoso De Almeida ML, Amaral De Castilho Valavicius;
 PI Gomes De Amorim Filho A;
 XX
 DR WPI; 1995-302720/39.
 DR N-PSDB; AA099460.
 XX
 PT Recombinant prodn. of proteins, e.g. insulin - by producing the protein
 PT with a glycosylphosphatidylinositol anchor followed by selective
 PT release.
 XX
 PS Disclosure; Fig 3; 51pp; English.
 XX
 CC Human recombinant insulin may be expressed in Saccharomyces cerevisiae
 CC following linkage of the gene to the glycosylphosphatidylinositol anchor.
 CC This anchoring technique can provide for the release of the product in a
 CC highly specific and selective manner. In addition, the recombinant
 CC protein will contain an epitope which can be used in its final
 CC purification by immunofluorescence. The protein product can be released by

CC e.g. nitrous deamination or treatment with neutral detergent. (Updated on
 CC 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 160 AA;
 Query Match 49.1%; Score 288.5; DB 2; Length 160;
 Best Local Similarity 96.1%; Pred. No. 2.5e-17; 0; Indels 1; Gaps 1;
 Matches 53; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 55 RFVNOHLGSHLVEALYLVGGERGFFYTPKTRGIVEQCCTSIQSLYLENYCN 107
 43 KFVNOHLGSHLVEALYLVGGERGFFYTPKTRGIVEQCCTSIQSLYLENYCN 96
 DB
 Search completed: November 2, 2004, 20:11:48
 Job time: 104.249 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:02:41 (Search time 26.059 Seconds
(without alignments)
272.306 Million cell updates/sec

Title: US-10-054-873-6

Perfect score: 587
Sequence: 1 MPTPLSRLEPDNMLRAHR.....IVQCTSLCYOLENYCN 107

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents PA:
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	304	51.8	63	US-08-160-376A-6	Sequence 6, Appl
2	302.5	51.5	137	US-08-400-256-38	Sequence 39, Appl
3	302.5	51.5	137	US-08-975-365-39	Sequence 39, Appl
4	299.5	51.0	66	US-08-291-060B-5	Sequence 5, Appl
5	299	50.9	56	US-08-160-376A-7	Sequence 7, Appl
6	299	50.9	56	US-08-389-487-11	Sequence 11, Appl
7	299	50.9	96	US-08-160-376A-5	Sequence 5, Appl
8	299	50.9	96	US-08-389-487-8	Sequence 8, Appl
9	299	50.9	145	US-08-400-256-45	Sequence 45, Appl
10	299	50.9	145	US-08-975-365-45	Sequence 45, Appl
11	299	50.9	146	US-08-400-256-48	Sequence 48, Appl
12	299	50.9	146	US-08-975-365-48	Sequence 48, Appl
13	299	49.9	57	US-08-030-731A-4	Sequence 44, Appl
14	287	48.9	65	US-08-900-574-3	Sequence 5, Appl
15	286.5	48.8	65	US-08-900-574-5	Sequence 5, Appl
16	286	48.7	67	US-08-900-574-7	Sequence 7, Appl
17	284.5	48.5	65	US-08-468-674B-71	Sequence 71, Appl
18	284.5	48.5	65	US-08-780-571-71	Sequence 71, Appl
19	284.5	48.5	124	US-09-012-669F-36	Sequence 36, Appl
20	284.5	48.5	124	US-09-894-711-18	Sequence 18, Appl
21	284	48.4	138	US-08-932-087-19	Sequence 19, Appl
22	284	48.4	138	US-09-861-687-19	Sequence 19, Appl
23	284	48.4	140	US-08-400-256-33	Sequence 33, Appl
24	284	48.4	140	US-08-400-256-42	Sequence 42, Appl
25	284	48.4	140	US-08-975-365-33	Sequence 33, Appl
26	284	48.4	140	US-08-975-365-42	Sequence 42, Appl
27	283.5	48.3	53	US-08-233-617-4	Sequence 4, Appl

28	283.5	48.3	53	US-08-981-988A-42	Sequence 42, Appl
29	283.5	48.3	117	US-09-012-669F-37	Sequence 37, Appl
30	281	47.9	104	US-08-400-256-15	Sequence 15, Appl
31	281	47.9	104	US-08-975-365-15	Sequence 15, Appl
32	280.5	47.8	89	US-08-468-674B-41	Sequence 41, Appl
33	280.5	47.8	89	US-08-780-571-41	Sequence 41, Appl
34	280.5	47.8	91	US-08-468-674B-45	Sequence 45, Appl
35	280.5	47.8	91	US-08-780-571-45	Sequence 45, Appl
36	280.5	47.8	124	US-08-446-646-3	Sequence 3, Appl
37	279.5	47.6	167	US-07-918-953-8	Sequence 8, Appl
38	279.5	47.6	167	US-08-081-661-8	Sequence 8, Appl
39	278.5	47.4	51	US-09-477-924-3	Sequence 3, Appl
40	278.5	47.4	51	US-09-723-981-3	Sequence 3, Appl
41	278.5	47.4	51	US-09-723-896-3	Sequence 3, Appl
42	278	47.4	117	US-08-280-030-63	Sequence 63, Appl
43	277.5	47.3	53	US-08-233-617-3	Sequence 3, Appl
44	277	47.2	96	US-09-134-836-4	Sequence 4, Appl
45	277	47.2	96	US-09-386-303A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-160-376A-6
Sequence 6, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Luckwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
OPERATING SYSTEM: WINDOWS 3.1
COMPUTER: IBM 386
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-6
Query Match 51.8%, Score 304, DB 1, Length 63;
Best Local Similarity 94.7%, Pred. No. 1,1e-28;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 51 GTGPRVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVGQCTSIQSLVQLENYCN 107
DB 7 GMSAFVNOHLCGSHLVEALYLVCGERGFFYTPKTRIGIVGQCTSIQSLVQLENYCN 63

RESULT 2

US-08-400-256-39
Sequence 39, Application US/08400256
Patent No. 5750497

GENERAL INFORMATION:

APPLICANT: Havelund, Svered
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Idn
APPLICANT: Andersen, Asger Sloth
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 57504970 No. 57504970 of No. 57504970th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor

CITY: New York
STATE: New York

COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-400-256-39

Query Match

Best Local Similarity 51.5%; Score 302.5; DB 1; Length 137;
Best Local Similarity 50.0%; Pred. No. 4.2e-28;
Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4;

QY 2 PPTPLSLRPFNMLRHRHLQALPDYQEFEEAYIPKQ--KYSFLQ-----N 48
DB 3 FPSI-----FPAVLPASSALAPVNTTDETAQIPAAVIVGSDLEDPVAVLPFSN 57

QY 49 PLGTG-----PRFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 87
DB 58 STNNGLFINTTASIAAKEGVSMARFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 117

QY 58 STNNGLFINTTASIAAKEGVSMARFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 117
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60110070 No. 60110070 of No. 60110070th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor

CITY: New York
STATE: New York

COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,365
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400,256
FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-975-365-39

Query Match 51.5%; Score 302.5; DB 3; Length 137;
Best Local Similarity 50.0%; Pred. No. 4.2e-28;
Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4;

QY 2 PPTPLSLRPFNMLRHRHLQALPDYQEFEEAYIPKQ--KYSFLQ-----N 48
DB 3 FPSI-----FPAVLPASSALAPVNTTDETAQIPAAVIVGSDLEDPVAVLPFSN 57

QY 49 PLGTG-----PRFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 87
DB 58 STNNGLFINTTASIAAKEGVSMARFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 117

QY 58 STNNGLFINTTASIAAKEGVSMARFVNOHLCGSHLVEALYLVCGERGFFYTPKTRG 117
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

QY 58 IVEQCTSIQSLVQLENYCN 107
DB 118 IVEQCTSIQSLVQLENYCN 137

Sat Nov 6 18:59:29 2004

us-10-054-873-6.ra1

Page 3

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231,060B
FILING DATE: 08-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Elnaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1105-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4366
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-291-060B-5

Query Match 51.0%; Score 299.5; DB 1; Length 66;
Best Local Similarity 91.7%; Pred. No. 3.9e-28;
Matches 55; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 48 RPLGTGPRFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 107
Db 8 DPNQNG-RFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 66

RESULT 5
US-08-160-376A-7
Sequence 7, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeyer, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
NUMBER OF INVENTION: Cysteine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:

TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-7

Query Match 50.9%; Score 299; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 RFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 107
Db 4 RFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 56

RESULT 6
US-08-389-487-11
Sequence 11, Application US/08389487
Patent No. 5663291
GENERAL INFORMATION:
APPLICANT: Obermeyer, Rainer
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jürgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having
TITLE OF INVENTION: Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dumer
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,487
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Elnaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1424-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-389-487-11

Query Match 50.9%; Score 299; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 RFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 107
Db 4 RFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIYVOCCTSIQSLYLQENYCN 56

RESULT 7

US-08-160-376A-5
Sequence 5, Application US/08160376A
Patent No. 5473049
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Proinsulin
TITLE OF INVENTION: Possessing Correctly Linked
TITLE OF INVENTION: Cysteine Bridges
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenneth A. Genoni, Esq.
STREET: Rt. 202-206 No. 5473049th/P.O. Box 2500
CITY: Somerville
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08876-1258
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 386
OPERATING SYSTEM: WINDOWS 3.1
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,376A
FILING DATE: December 1, 1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GE P 4240420.7
FILING DATE: December 2, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Barbara V. Maurer, Esq.
REGISTRATION NUMBER: 31,287
REFERENCE/DOCKET NUMBER: HOE 92/F 384
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 231-4079
TELEFAX: (908) 231-2255
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 Amino Acids
TYPE: Amino Acid (AA)
TOPOLOGY: not relevant
US-08-160-376A-5

Query Match 50.9%; Score 299; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 7,1e-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 107
DB 44 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 96

RESULT 8
US-08-389-487-8
Sequence 8, Application US/08389487
Patent No. 5663291
GENERAL INFORMATION:
APPLICANT: Obermeier, Ranier
APPLICANT: Gerl, Martin
APPLICANT: Ludwig, Jurgen
APPLICANT: Sabel, Walter
TITLE OF INVENTION: Process for Obtaining Insulin Having
TITLE OF INVENTION: Correctly Linked Cysteine Bridges
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States of America

ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/389,487
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1424-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-389-487-8

Query Match 50.9%; Score 299; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 7,1e-28;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 107
DB 44 RFVNHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIQSLYQLENYCN 96

RESULT 9
US-08-400-266-45
Sequence 45, Application US/08400266
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,266
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-400-256-45

Query Match 50.9%; Score 299; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 107
DB 93 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 145

RESULT 10

US-08-975-365-45
Sequence 45, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975.365
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400.256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-975-365-45

Query Match 50.9%; Score 299; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 107
DB 93 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 145

RESULT 11

US-08-400-256-48
Sequence 48, Application US/08400256
Patent No. 5750497

GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 57504970 No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

QY 55 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 107
DB 94 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 146

Query Match 50.9%; Score 299; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 107
DB 94 RFVNOHLCGSHLVEALYIVCGERGFYPTKTRGIVEQCTISCSLYOLENYCN 146

RESULT 12

US-08-975-365-48
Sequence 48, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/975,365
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-975-365-48

Query Match 50.9%; Score 299; DB 3; Length 146;
Best Local Similarity 100.0%; Pred. No. 1,2e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

55 RPNVHLCGSHLVETLVYVCGERGFYTPKTRGIVEQCCTSLCYOLENYCN 107
Db 94 RPNVHLCGSHLVETLVYVCGERGFYTPKTRGIVEQCCTSLCYOLENYCN 146

RESULT 13
US-08-030-731A-44
Sequence 44, Application US/08030731A
Patent No. 5426036
GENERAL INFORMATION:
APPLICANT: Koller, Klaus-Peter
APPLICANT: Riess, Guenther Johannes
APPLICANT: Uhlmann, Eugen
APPLICANT: Walmeier, Holger
TITLE OF INVENTION: Processes for the Preparation of Foreign
TITLE OF INVENTION: Proteins in Streptomyces
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,731A
FILING DATE: 12-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/189,840
FILING DATE: 03-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/430,622
FILING DATE: 01-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/687,610
FILING DATE: 19-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,757
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 37 14 866.4

FILING DATE: 05-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 38 37 273.8
FILING DATE: 03-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 35 27 449.7
FILING DATE: 19-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 40 12 818.0
FILING DATE: 21-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Kirschner, Michael K.
REGISTRATION NUMBER: 34,851
REFERENCE/DOCKET NUMBER: 02481-0593-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-030-731A-44

Query Match 49.9%; Score 293; DB 1; Length 57;
Best Local Similarity 96.2%; Pred. No. 1,9e-27;
Matches 51; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

55 RPNVHLCGSHLVETLVYVCGERGFYTPKTRGIVEQCCTSLCYOLENYCN 107
Db 5 RPNVHLCGSHLVETLVYVCGERGFYTPKTRGIVEQCCTSLCYOLENYCN 57

RESULT 14
US-08-900-574-3
Sequence 3, Application US/08900574
Patent No. 6221837
GENERAL INFORMATION:
APPLICANT: Ertl, Johann
APPLICANT: Habermann, Paul
APPLICANT: Geispe, Karl
TITLE OF INVENTION: Insulin derivatives with increased zinc
TITLE OF INVENTION: binding
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett,
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,574
FILING DATE: July 24, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: German Application No. 6221837 19630242.0
FILING DATE: July 26, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Carol P. Einaudi
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481,1499-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: Protein
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Protein
LOCATION: 1..65
US-08-900-574-3

Query Match 48.9%; Score 287; DB 3; Length 65;
Best Local Similarity 91.4%; Pred. No. 1.1e-26;
Matches 53; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 51 GTGPRFVNQHLGSHLVEALYLVCGERGFFYTPKT--RGIYEOCCCTSIICSLYQLENYC 106
DB 7 GNSARFVNQHLGSHLVEALYLVCGERGFFYTPKTHRHGIVGECCTSIICSLYQLENYC 64

RESULT 15

US-08-900-574-5
Sequence 5, Application US/08900574
Patent No. 6221837
GENERAL INFORMATION:
APPLICANT: Ertl, Johann
APPLICANT: Habermann, Paul
APPLICANT: Geisen, Karl
APPLICANT: Seipke, Gerhard
TITLE OF INVENTION: Insulin derivatives with increased zinc
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett,
ADDRESSEE: & Dunne, L.L.P.
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: U.S.A.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,574
FILING DATE: July 24, 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: German Application No. 6221837 19630242.0
FILING DATE: July 26, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Carol P. Einaudi
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481.1499-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: linear
MOLECULE TYPE: Protein
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: Protein

LOCATION: 1..66
US-08-900-574-5

Query Match 48.8%; Score 286.5; DB 3; Length 66;
Best Local Similarity 89.8%; Pred. No. 1.3e-26;
Matches 53; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 51 GTGPRFVNQHLGSHLVEALYLVCGERGFFYTPKT--RGIYEOCCCTSIICSLYQLENYC 106
DB 7 GNSARFVNQHLGSHLVEALYLVCGERGFFYTPKTHRHGIVGECCTSIICSLYQLENYC 65

Search completed: November 2, 2004, 20:24:35
Job time: 26.059 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:59:41 : Search time 19.7417 seconds

(without alignments)
521.495 Million cell updates/sec

Title: US-10-054-873-6

Perfect score: 587

Sequence: 1 MFPTPLSLRPLFDNMLRAHR.....IVEQCTSTCSISVLENYCN 107

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: *
2: Dair2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	275	46.8	96	2	PC7082
2	273.5	46.6	51	1	INEL
3	273.5	46.6	51	1	INMHP
4	273.5	46.6	51	1	INMHP
5	273	46.5	110	2	B42179
6	273	46.5	110	2	U00178
7	271.5	46.3	51	1	INNY
8	268.5	45.7	51	1	INMSP
9	267.5	45.6	51	2	A59151
10	267	45.5	110	2	A42179
11	267	45.5	110	2	A42179
12	263.5	44.9	51	1	INMCA
13	263.5	44.9	51	1	INMT
14	263.5	44.9	51	1	INMHS
15	263	44.8	84	1	IPG
16	262.5	44.7	110	1	INNB
17	262.5	44.7	51	1	INCT
18	262	44.6	110	1	IPG
19	261.5	44.5	51	1	INMKS
20	260	44.3	110	2	I48166
21	258.5	44.0	105	1	I180
22	257	43.8	108	2	A39883
23	256.5	43.7	51	2	U00362
24	255.5	43.5	217	1	STW
25	255.5	43.5	217	1	I67410
26	252.5	43.0	77	1	INSH
27	252	42.9	86	1	IPHO
28	251.5	42.8	51	1	INCB
29	250	42.6	108	1	INMS1

30	249	42.4	110	1	IPRT1	insulin 1 precursor
31	248.5	42.3	51	1	INGS	insulin - goose
32	248	42.2	110	1	INMS2	insulin 2 precursor
33	248	42.2	110	1	IPRT2	insulin 2 precursor
34	246	41.9	52	2	S44470	insulin 12 - North
35	246	41.9	52	2	S44469	insulin 11 - North
36	245	41.7	103	2	I51221	insulin precursor
37	244.5	41.7	51	1	INOS	insulin - ostrich
38	244.5	41.7	51	1	INTK	insulin - turkey
39	244.5	41.7	51	1	INP129	insulin - black-be
40	244.5	41.7	51	1	INPO	insulin - crested
41	244.5	41.7	51	2	A60414	insulin - slider t
42	239.5	40.8	107	1	IPCH	insulin precursor
43	238	40.5	52	2	S61361	insulin - Amphiuma
44	235.5	40.1	51	2	S61350	insulin - duckbill
45	233.5	39.8	81	1	IPDK	insulin precursor

ALIGNMENTS

RESULT 1

PC7082

epidermal growth factor/single chain insulin fusion protein - Bacillus brevis (fragmen

C/Species: Bacillus brevis

C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004

C/Accession: PC7082; PC7083

R/Koh: M.; Hanagata, H.; Ebisu, S.; Morihara, K.; Takagi, H.

Biosci. Biotechnol. Biochem. 64, 1079-1081, 2000

A/Title: Use of Bacillus brevis for synthesis and secretion of Des-B30 single-chain hu

A/Reference number: PC7082; MUID:20335834; PMID:10879487

A/Accession: PC7082

A/Molecule type: DNA

A/Residues: 1-96 <KOH>

A/Cross-references: UNIPROT:Q7M0U6

A/Accession: PC7083

A/Molecule type: protein

A/Residues: 19-28 <K02>

C/Genetics:

A/Gene: egf-8cl

C/Superfamily: Insulin

Query Match

Best Local Similarity 46.8%; Score 275; DB 2; Length 96;

Matches 50; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

DB 55 RFVNOHLCGSHVAVLVGCEGFPTPTKRGIVEQCTSTCSISVLENYCN 107

46 RFVNOHLCGSHVAVLVGCEGFPTPTKRGIVEQCTSTCSISVLENYCN 96

RESULT 2

INEL

insulin - elephant

C/Species: Elephantidae gen. sp. (elephant)

C/Date: 24-Apr-1984 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1993

C/Accession: A01584

R/Smith, L.F.

Am. J. Med. 40, 662-666, 1966

A/Title: Species variation in the amino acid sequence of insulin.

A/Reference number: A00029; MUID:66160119; PMID:5949593

A/Accession: A01584

A/Molecule type: protein

A/Residues: 1-30;31-51 <GMT>

A/Note: The species of elephant is not given, but it is most probably the Indian eleph

C/Superfamily: Insulin

C/Keywords: hormone; pancreas

F/1-30/Domain: Insulin chain B #status experimental <BCH>

F/1-30/31-51/Product: Insulin #status experimental <MAT>

F/31-51/Domain: Insulin chain A #status experimental <ACH>

F/7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match

46.6%; Score 273.5; DB 1; Length 51;

Best Local Similarity 94.2%; Pred. No. 1.7e-21;
Matches 49; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 107

1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 51

RESULT 3

INMHF

Insulin - finback whale (tentative sequence)

C:Species: Baleenoptera physalus (finback whale, common rorqual)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: A91918

R:Hama, H.; Tizabi, K.; Sakaki, S.; Narita, K.

U:Biochem. 56, 285-293, 1964

A:Title: The amino acid sequence in fin-whale insulin.

A:Reference number: A91918

A:Accession: A91918

A:Molecule type: Protein

A:Residues: 1-30;31-51 <HAM>

A:Cross-references: UNIPROT:P01312

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-30/Domain: Insulin chain B #status experimental <BCH>

F:1-30,31-51/Product: Insulin #status experimental <MAT>

F:31-51/Domain: Insulin chain A #status experimental <ACH>

F:7-37,49-50,36-41/Dissulfide bonds: #status predicted

Query Match 46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 1.7e-21;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 107

1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 51

RESULT 4

INMHP

Insulin - sperm whale

C:Species: Physeter catodon (sperm whale)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: A93142; A90082

R:Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.

U:Nature 181, 1468-1469, 1958

A:Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pepsin.

A:Reference number: A93142

A:Accession: A93142

A:Molecule type: Protein

A:Residues: 1-30;31-51 <ISH>

A:Cross-references: UNIPROT:P01312

R:Harris, J.I.; Sanger, F.; Naughton, M.A.

U:Arch. Biochem. Biophys. 65, 427-428, 1956

A:Title: Species differences in insulin.

A:Reference number: A90082

A:Accession: A90082

A:Molecule type: Protein

A:Residues: 1-30;31-51 <HAR>

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-30/Domain: Insulin chain B #status experimental <BCH>

F:1-30,31-51/Product: Insulin #status experimental <MAT>

F:31-51/Domain: Insulin chain A #status experimental <ACH>

F:7-37,49-50,36-41/Dissulfide bonds: #status predicted

Query Match 46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 1.7e-21;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 107

1 FVNQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 51

RESULT 5

Insulin precursor - green monkey

C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: B42179; A05232; S16494; S22056

R:Seino, S.; Bell, G.I.; Li, W.H.

U:Mol. Biol. Evol. 9, 193-203, 1992

A:Title: Sequences of primate insulin genes support the hypothesis of a slower rate of

A:Reference number: A42179; PMID:92219953; PMID:1560757

A:Accession: B42179

A:Molecule type: DNA

A:Residues: 1-110 <SEI>

A:Cross-references: UNIPROT:P30407; EMBL:X61092; NID:G22808; PIDN:CAA43405.1; PID:G2280

A>Note: Sequence extracted from NCBI backbone (NCBIN:95185; NCBI:95194)

R:Peterson, U.D.; Neirlich, S.; Oyer, F.E.; Steiner, D.F.

U: J. Biol. Chem. 247, 4866-4871, 1972

A:Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsu

A:Reference number: A92111; PMID:72258016; PMID:4626369

A:Accession: A05232

A:Molecule type: Protein

A:Residues: 57-87 <PET>

C:Genetics: 63/1

C:Superfamily: Insulin

C:Keywords: hormone; pancreas

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-54/Domain: insulin chain B #status predicted <BCH>

F:25-54,90-110/Product: Insulin #status predicted <MAT>

F:57-87/Domain: connecting peptide #status experimental <CEP>

F:90-110/Domain: insulin chain A #status predicted <ACH>

F:31-96,43-109,95-100/Dissulfide bonds: #status predicted

Query Match 46.5%; Score 273; DB 2; Length 110;
Best Local Similarity 60.2%; Pred. No. 4.2e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PRFVQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 85

23 PRFVQHLGSHLVEALYLVCGERGFFYPTKTRGIVEQCCTSIGVQLVLENYCN 82

QY 86 -----RGIVEQCCTSIGVQLVLENYCN 107

Db 83 EGSLOQRGIVEQCCTSIGVQLVLENYCN 110

RESULT 6

Insulin precursor - crab-eating macaque

C:Species: Macaca fascicularis (crab-eating macaque)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004

C:Accession: J00178

R:Metzke, W.; Gronenberg, J.; Leineweber, M.; Wengenmayer, F.; Wimmacker, E.L.

U:Gene 19, 179-183, 1982

A:Title: The nucleotide sequence of cDNA coding for proproinsulin from the primate Mac

A:Reference number: J00178; PMID:83080474; PMID:6184262

A:Accession: J00178

A:Molecule type: mRNA

A:Residues: 1-110 <MET>

A:Cross-references: UNIPROT:P30406; GB:U00336; NID:G342121; PIDN:AAA36849.1; PID:G3421

C:Superfamily: Insulin

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-54,90-110/Product: Insulin #status predicted <MAT>

F:25-54/Domain: insulin chain B #status predicted <BCH>

F:55-89/Domain: insulin connecting C peptide #status predicted <CEP>

F:90-110/Domain: insulin chain A #status predicted <ACH>

F:31-96,43-109,95-100/Dissulfide bonds: #status predicted

Query Match 46.5%; Score 273; DB 2; Length 110;
Best Local Similarity 60.2%; Pred. No. 4.2e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PRVNOHLCGSHLVEALYVCGERGFYTPKT----- 85
 Db 23 PARVNOHLCGSHLVEALYVCGERGFYTPKTRREADPOVQVELGGGPGAGSLPIAL 82
 QY 86 -----RGIVEOCCTSCSLYOLENYCN 107
 Db 83 EGSIGRGIVGOCCTSCSLYOLENYCN 110

RESULT 7
 INHY
 Insulin - hamster
 C:Species: Cricetinae gen. sp. (hamster)
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: A91456
 R:Neelon, F.A.; Delcher, H.K.; Steinman, H.; Lebovitz, H.E.
 Fed. Proc. 32, 300, 1973
 A:Title: Structure of hamster insulin: comparison with a tumor insulin.
 A:Reference number: A91456
 A:Accession: A91456
 A:Molecule type: protein
 A:Residues: 1-30/31-51 <NE>
 A:Cross-references: UNIPROT:Q7M0G1
 C:Superfamily: Insulin
 F:1-30/Domain: Insulin chain B #status experimental <BCH>
 F:1-30/31-51/Product: Insulin #status experimental <MAT>
 F:31-51/Domain: Insulin chain A #status experimental <ACH>
 F:7-37/19-50/36-41/Disulfide bonds: #status predicted

Query Match 46.3%; Score 271.5; DB 1; Length 51;
 Best Local Similarity 94.2%; Pred. No. 2,7e-21;
 Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYVCGERGFYTPKTRGIVEOCCTSCSLYOLENYCN 107
 Db 1 FVNQHLGSHLVEALYVCGERGFYTPKS-GIVDQCTSCSLYOLENYCN 51

RESULT 8
 INMSSP
 Insulin - Egyptian spiny mouse (tentative sequence)
 C:Species: Acromys caninus (Egyptian spiny mouse)
 C>Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
 C:Accession: A01591
 R:Brizell, H.F.; Humbel, R.E.
 Hope-Seyler's Z. Physiol. Chem. 353, 444-450, 1972
 A:Title: Isolation and partial structural analysis of insulin from mouse (Mus musculus)
 A:Reference number: A01591; PMID:72189454; PMID:5028210
 A:Contents: composition
 A:Accession: A01591
 A:Molecule type: protein
 A:Residues: 1-30/31-51 <BDE>
 A:Cross-references: UNIPROT:P01324
 C:Superfamily: Insulin
 C:Keywords: hormone; pancreas
 F:1-30/Domain: Insulin chain B #status predicted <BCH>
 F:1-30/31-51/Product: Insulin #status predicted <MAT>
 F:31-51/Domain: Insulin chain A #status predicted <ACH>
 F:7-37/19-50/36-41/Disulfide bonds: #status predicted

Query Match 45.7%; Score 268.5; DB 1; Length 51;
 Best Local Similarity 92.3%; Pred. No. 5,5e-21;
 Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYVCGERGFYTPKTRGIVEOCCTSCSLYOLENYCN 107
 Db 1 FVBGHLGSHLVEALYVCGERGFYTPKS-GIVDQCTSCSLYOLENYCN 51

RESULT 9
 A59151

Insulin precursor - jack bean (fragments)
 N:Alternate names: hypoglycemic agent; plant insulin
 C:Species: Canavalia ensiformis (jack bean)
 C>Date: 07-Dec-1999 #sequence_revision 07-Dec-1999 #text_change 10-Dec-1999
 C:Accession: B59151; A59151
 R:Oliveira, A.E.A.; Machado, O.L.T.; Gomes, V.M.; Xavier-Neto, J.; Pereira, A.C.F.; V
 Protein. Pept. Lett. 6, 15-21, 1999
 A:Title: Jack bean seed coat contains a protein with complete sequence homology to bo
 A:Reference number: A59151
 A:Accession: B59151
 A:Molecule type: protein
 A:Residues: 1-30 <MACB>
 A:Cross-references: UNIPROT:Q7M217
 A:Accession: A59151
 A:Molecule type: protein
 A:Residues: 31-51 <MACA>
 C:Comment: The two chains are probably produced from the same precursor.
 C:Superfamily: Insulin
 F:1-30/31-51/Product: Insulin #status experimental <MAT>
 F:1-30/Domain: chain B #status experimental <CHB>
 F:31-51/Domain: chain A #status experimental <CHA>
 F:7-37/19-50/36-41/Disulfide bonds: #status predicted

Query Match 45.6%; Score 267.5; DB 2; Length 51;
 Best Local Similarity 92.3%; Pred. No. 7e-21;
 Matches 48; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYVCGERGFYTPKTRGIVEOCCTSCSLYOLENYCN 107
 Db 1 FVNQHLGSHLVEALYVCGERGFYTPKA-GIVEOCASVCSLYOLENYCN 51

RESULT 10
 INHU
 Insulin precursor (validated) - human
 N:Alternate names: preproinsulin
 C:Species: Homo sapiens (man)
 C>Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 09-Jul-2004
 C:Accession: A93222; A94253; A93216; A93144; A92075; A91166; I58114; A01579; S
 R:Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischner, E.; Goodman, H.M.
 Nature 284, 26-32, 1980
 A:Title: Sequence of the human insulin gene.
 A:Reference number: A93222; PMID:80120725; PMID:6243748
 A:Accession: A93222
 A:Molecule type: DNA
 A:Residues: 1-110 <BEU>
 A:Cross-references: UNIPROT:P01308; GB:U00265; NID:G186429; PIDN:AAA59172.1; PID:G38682
 R:Ullrich, A.; Dull, T.J.; Gray, A.; Brosius, J.; Sures, I.
 Science 209, 612-615, 1980
 A:Title: Genetic variation in the human insulin gene.
 A:Reference number: A94253; PMID:80236313; PMID:6248962
 A:Accession: A94253
 A:Molecule type: DNA
 A:Residues: 1-110 <BEU>
 A:Cross-references: GB:U00265; NID:G186429; PIDN:AAA59172.1; PID:G38682
 R:Bell, G.I.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.J.
 Nature 282, 525-527, 1979
 A:Title: Nucleotide sequence of a cDNA clone encoding human preproinsulin.
 A:Reference number: A93216; PMID:80054779; PMID:503224
 A:Accession: A93216
 A:Molecule type: mRNA
 A:Residues: 1-110 <BEU>
 A:Cross-references: GB:U00265; NID:G186429; PIDN:AAA59172.1; PID:G38682
 R:Sures, I.; Goeddel, D.V.; Gray, A.; Ullrich, A.
 Science 208, 57-59, 1980
 A:Title: Nucleotide sequence of human preproinsulin complementary DNA.
 A:Reference number: A94251; PMID:80147417; PMID:6927840
 A:Accession: A94251
 A:Molecule type: mRNA
 A:Residues: 1-110 <SUP>
 A:Cross-references: GB:U00265; NID:G186429; PIDN:AAA59172.1; PID:G38682
 R:Nicol, D.S.H.W.; Smith, L.F.
 Nature 187, 483-485, 1960

A>Title: Amino-acid sequence of human insulin.
 A/Reference number: A93144
 A/Accession: A93144
 A/Molecule type: Protein
 A/Residues: 25-54;90-110 <NIC>
 R/Oyer, P.E.; Cho, S.; Peterson, J.D.; Steiner, D.F.
 J. Biol. Chem. 246, 1375-1386, 1971
 A>Title: Studies on human proinsulin. Isolation and amino acid sequence of the human part
 A/Reference number: A92075; MUID:71116410; PMID:5101771
 A/Accession: A92075
 A/Molecule type: protein
 A/Residues: 57-87 <OYE>
 R/Oyer, P.E.; Steiner, D.F.; Markussen, J.; Sundby, F.
 Eur. J. Biochem. 20, 190-199, 1971
 A>Title: Amino acid sequence of the C-peptide of human proinsulin.
 A/Reference number: A91186; MUID:71257722; PMID:5560404
 A/Accession: A91186
 A/Molecule type: protein
 A/Residues: 57-87 <KOA>
 R/ucassen, A.M.; Jøller, C.; Beressi, J.P.; Boltard, C.; Froguel, P.; Lathrop, M.; Bell
 Nature Genet. 4, 305-310, 1993
 A>Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
 A/Reference number: 158114; MUID:93364428; PMID:8358440
 A/Accession: 158114
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-59;63-110 <RES>
 A/Cross-references: GB:115440; NID:9307071; PIDN:AA59179.1; PID:9307072
 R/Sieber, P.; Kamber, B.; Hartmann, A.; Jöchl, A.; Rühner, B.; Rüttel, W.
 Helv. Chim. Acta 57, 2617-2621, 1974
 A>Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
 A/Reference number: A91636; MUID:75077277; PMID:4443293
 A/Accession: A91636
 A/Status: annotation; synthesis
 A/Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identical
 A/Note: article in German with English abstract
 R/Naehd, V.X.
 Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
 A>Title: The synthesis of C-peptide of human proinsulin.
 A/Reference number: A91658; MUID:75040007; PMID:4803504
 A/Accession: A91658
 A/Status: annotation; synthesis of residues 57-87
 R/Geiger, R.; Jaeger, G.; Koenig, W.
 Chem. Ber. 106, 2347-2352, 1973
 A>Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu-6
 A/Reference number: A90914
 A/Accession: A90914
 A/Status: annotation; synthesis of residues 57-87
 R/Baummann, J.E.; Irminger, J.C.; Halban, P.A.
 Biochem. J. 310, 869-874, 1995
 A>Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junction
 A/Reference number: S58651; MUID:96013185; PMID:7575420
 A/Accession: S58651
 A/Status: annotation; site-directed mutagenesis study of proteolytic processing
 C/Genetics:
 A/Genes: GDB:INS
 A/Cross-references: GDB:119349; OMIM:176730
 A/Map position: 11p15.5-11p15.5
 A/Introns: 63/1
 C/Keywords: hormone; pancreas
 C/Superfamily: insulin
 F/1-24/Domain: signal sequence #status predicted <SIG>
 F/25-54/Domain: insulin chain B #status experimental <BCH>
 F/57-87/Domain: insulin chain A #status experimental <MAT>
 F/90-110/Domain: connecting C peptide #status experimental <CPBP>
 F/31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 45.5%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 1.8e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQVGLGPGASGQSPLALEG 84
 QY 86 ----RGIVEQCCTSIICSLYQLENYCN 107

DB 85 SLKRGIVEQCCTSIICSLYQLENYCN 110

RESULT 11
 A42179
 Insulin precursor - chimpanzee
 C/Species: Pan troglodytes (chimpanzee)
 C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C/Accession: A42179; S22058
 R/Setno, S.; Bell, G.I.; Li, W.H.
 Mol Biol Evol. 9, 193-203, 1992
 A>Title: Sequences of primate insulin genes support the hypothesis of a slower rate of
 A/Reference number: A42179; MUID:92219553; PMID:1560757
 A/Accession: A42179
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-110 <SEI>
 A/Cross-references: UNIPROT:P30410; EMBL:X61089; NID:938251; PIDN:CAA43403.1; PID:938252
 A/Note: sequence extracted from NCBI backbone (NCBI:P30410)
 C/Genetics:
 A/Introns: 63/1
 C/Superfamily: insulin

Query Match 45.5%; Score 267; DB 2; Length 110;
 Best Local Similarity 60.5%; Pred. No. 1.8e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREAEDLQVQLQVGLGPGASGQSPLALEG 84
 QY 86 ----RGIVEQCCTSIICSLYQLENYCN 107
 DB 85 SLKRGIVEQCCTSIICSLYQLENYCN 110

RESULT 12
 INCMN
 Insulin - Arabian camel (tentative sequence)
 C/Species: Camelus dromedarius (Arabian camel)
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
 C/Accession: A92782
 R/Danho, W.O.
 J. Fac. Med. Baghdad 14, 16-28, 1972
 A>Title: The isolation and characterization of insulin of camel (Camelus dromedarius).
 A/Reference number: A92782
 A/Accession: A92782
 A/Molecule type: protein
 A/Residues: 1-30;31-51 <DAN>
 A/Cross-references: UNIPROT:P01320
 C/Superfamily: insulin
 C/Keywords: hormone; pancreas
 F/1-30/Domain: insulin chain B #status experimental <BCH>
 F/31-51/Domain: insulin chain A #status experimental <MAT>
 F/31-51/Domain: insulin chain A #status experimental <ACH>
 F/7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 44.9%; Score 263.5; DB 1; Length 51;
 Best Local Similarity 90.4%; Pred. No. 1.8e-20;
 Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIICSLYQLENYCN 107
 DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSIICSLYQLENYCN 51

RESULT 13
 INGT
 Insulin - goat
 C/Species: Capra aegagrus hircus (domestic goat)
 C/Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
 C/Accession: A01586

R:Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; PMID:66160119; PMID:5949593
A:Accession: A01586
A:Molecule type: protein
A:Residues: 1-30:31-51 <SMT>
A:Cross-references: UNIPROT:P01319
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: Insulin chain B #status experimental <BCH>
F:1-30,31-51/Product: Insulin #status experimental <MAT>
F:31-51/Domain: Insulin chain A #status experimental <ACH>
F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 90.4%; Pred. No. 1.8e-20;
Matches 47; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRIGIVQCCTSIQSLYLENYCN 107
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCAGVCSLYOLENYCN 51

RESULT 14

INMH15
Insulin - sei whale
C:Species: Balaeoptera borealis (sei whale)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1991 #text_change 09-Jul-2004
C:Accession: A01582
R:Ichihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 181, 1468-1469, 1956
A:Title: Structure of sperm- and sei-whale insulins and their breakdown by whale pepsin.
A:Reference number: A93142
A:Accession: A01582
A:Molecule type: protein
A:Residues: 1-30:31-51 <ISH>
A:Cross-references: UNIPROT:P01314
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-30/Domain: Insulin chain B #status experimental <BCH>
F:1-30,31-51/Product: Insulin #status experimental <MAT>
F:31-51/Domain: Insulin chain A #status experimental <ACH>
F:7-37,19-50,36-41/Disulfide bonds: #status predicted

Query Match 44.9%; Score 263.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 1.8e-20;
Matches 48; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRIGIVQCCTSIQSLYLENYCN 107
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKA-GIVEQCAGVCSLYOLENYCN 51

RESULT 15

IPG

Insulin precursor - pig

C:Species: Sus scrofa domestica (domestic pig)
C:Date: 22-Jun-1981 #sequence_revision 22-Jun-1991 #text_change 16-Jul-1999
C:Accession: A01583; A94572; S16492; A60835; B60835

R:Chance, R.E.; Ellis, R.W.; Bromer, W.W.
Science 161, 165-167, 1968

A:Title: Porcine proinsulin: characterization and amino acid sequence.
A:Reference number: A94240; PMID:68286485; PMID:5657063
A:Accession: A01583

A:Molecule type: protein
A:Residues: 1-34; 'Q', 36-84 <CHA>

R:Chance, R.E.
submitted to the Atlas, July 1970

A:Reference number: A94572
A:Accession: A94572

A:Molecule type: protein
A:Residues: 1-84 <CH2>

R:Brown, H.; Sanger, F.; Kitai, R.
Biochem. J. 60, 556-565, 1955
A:Title: The structure of pig and sheep insulins.
A:Reference number: A90344
A:Accession: S16492

A:Molecule type: protein
A:Residues: 1-30:31-51<BRO>

R:Snell, L.; Damgaard, U.
Horm. Metab. Res. 20, 476-480, 1988

A:Title: Proinsulin heterogeneity in pigs.
A:Reference number: A60835; PMID:89032178; PMID:3181865
A:Accession: A60835

A:Molecule type: protein
A:Residues: 33-38,40-62 <SNB>

A:Note: the authors report the characterization of a connecting peptide variant lacki

A:Accession: B60835
A:Molecule type: protein

A:Residues: 33-62 <SN2>
R:Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D.
Adv. Protein Chem. 26, 279-402, 1972

A:Title: Insulin: the structure in the crystal and its reflection in chemistry and bio

A:Reference number: A90017
A:Contents: annotation; X-ray crystallography, 1.9 angstroms
C:Superfamily: insulin

C:Keywords: hormone; pancreas
F:1-30/Domain: Insulin chain B #status experimental <BCH>
F:1-30,64-84/Product: Insulin #status experimental <MAT>

F:33-63,64-84/Product: connecting peptide #status experimental <CEPP>
F:64-84/Domain: Insulin chain A #status experimental <ACH>

F:7-70,19-83,69-74/Disulfide bonds: #status experimental

Query Match 44.8%; Score 263; DB 1; Length 84;
Best Local Similarity 60.7%; Pred. No. 3.4e-20;
Matches 51; Conservative 0; Mismatches 1; Indels 32; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYPTKTRIGIVQCCTSIQSLYLENYCN 85
DB 1 FVNQHLGSHLVEALYLVCGERGFFYTPKARRRREANPDGAVELGGGLGLOALALEGPP 60

QY 86 --RGIVQCCTSIQSLYLENYCN 107
DB 61 QKRGIVQCCTSIQSLYLENYCN 84

Search completed: November 2, 2004, 20:22:16
Job time: 19.7417 secs

This Page Blank (uspto)

Sat Nov 6 18:59:29 2004

us-10-054-873-6.rabb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:20:47; Search time 80.1513 Seconds
(without alignments)
432.820 Million cell updates/sec

Title: US-10-054-873-6

Sequence: 1 MFPTPLSRPLFDNMLRAHR.....IVEGCTSLGSLVQLNYCN 107

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	Published Applications AA:*
1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
3:	/cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
4:	/cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep:*
6:	/cgn2_6/ptodata/1/pubpaa/US02_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/1/pubpaa/US01_PUBCOMB.pep:*
8:	/cgn2_6/ptodata/1/pubpaa/US00_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
13:	/cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/1/pubpaa/US04_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/1/pubpaa/US03_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/1/pubpaa/US02_PUBCOMB.pep:*
17:	/cgn2_6/ptodata/1/pubpaa/US01_PUBCOMB.pep:*
18:	/cgn2_6/ptodata/1/pubpaa/US00_PUBCOMB.pep:*
19:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
20:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	587	100.0	107	US-10-054-873-6	Sequence 6, Appl1
2	555.5	94.6	150	US-10-054-873-7	Sequence 7, Appl1
3	302.5	51.5	137	US-10-101-454-39	Sequence 39, Appl1
4	299	50.9	145	US-10-101-454-45	Sequence 45, Appl1
5	294	50.1	146	US-10-101-454-48	Sequence 48, Appl1
6	294	50.1	152	US-10-054-873-5	Sequence 5, Appl1
7	284.5	48.5	124	US-09-894-711-18	Sequence 18, Appl1
8	284	48.4	138	US-09-861-687-19	Sequence 19, Appl1
9	284	48.4	138	US-10-620-651-19	Sequence 19, Appl1
10	284	48.4	140	US-10-101-454-33	Sequence 33, Appl1
11	284	48.4	140	US-10-101-454-42	Sequence 42, Appl1
12	281	47.9	104	US-10-101-454-15	Sequence 15, Appl1
13	278.5	47.4	51	US-09-858-935B-5	Sequence 5, Appl1

14	278.5	47.4	51	US-10-028-410-3	Sequence 3, Appl1
15	278.5	47.4	51	US-10-444-326-3	Sequence 3, Appl1
16	278.5	47.4	51	US-10-271-869-5	Sequence 5, Appl1
17	278.5	47.4	51	US-10-444-262-3	Sequence 3, Appl1
18	278.5	47.4	51	US-10-444-649-3	Sequence 3, Appl1
19	278.5	47.4	51	US-10-444-701-3	Sequence 3, Appl1
20	278.5	47.4	51	US-09-280-030-63	Sequence 63, Appl1
21	278.5	47.3	124	US-10-221-677-24	Sequence 24, Appl1
22	277	47.2	96	US-09-947-563-4	Sequence 4, Appl1
23	277	47.2	102	US-10-101-454-36	Sequence 36, Appl1
24	275.5	46.9	124	US-09-736-611-12	Sequence 12, Appl1
25	275.5	46.9	124	US-09-740-359-12	Sequence 12, Appl1
26	275.5	46.9	124	US-09-694-711-12	Sequence 12, Appl1
27	275.5	46.9	124	US-10-316-421-12	Sequence 12, Appl1
28	275.5	46.9	125	US-09-736-611-10	Sequence 10, Appl1
29	275.5	46.9	125	US-09-740-359-10	Sequence 10, Appl1
30	275.5	46.9	125	US-09-694-711-10	Sequence 10, Appl1
31	275.5	46.9	125	US-10-316-421-10	Sequence 10, Appl1
32	275.5	46.9	147	US-09-736-611-8	Sequence 8, Appl1
33	275.5	46.9	147	US-09-740-359-7	Sequence 7, Appl1
34	275.5	46.9	147	US-10-316-421-8	Sequence 8, Appl1
35	274	46.7	144	US-09-736-611-6	Sequence 6, Appl1
36	274	46.7	144	US-09-740-359-5	Sequence 5, Appl1
37	274	46.7	144	US-10-316-421-6	Sequence 6, Appl1
38	274	46.7	146	US-09-894-711-5	Sequence 5, Appl1
39	273	46.5	50	US-10-066-003A-3	Sequence 3, Appl1
40	271	46.2	96	US-09-947-563-5	Sequence 5, Appl1
41	270	46.0	104	US-10-101-454-21	Sequence 21, Appl1
42	270	46.0	104	US-10-101-454-27	Sequence 27, Appl1
43	269.5	45.9	130	US-09-280-030-62	Sequence 62, Appl1
44	269	45.8	104	US-10-101-454-24	Sequence 24, Appl1
45	269	45.8	104	US-10-101-454-30	Sequence 30, Appl1

ALIGNMENTS

RESULT 1
US-10-054-873-6
Sequence 6, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-054-873-6

Query Match 100.0%; Score 587; DB 13; Length 107;
Best Local Similarity 100.0%; Pred. No. 7, 7e-61;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 60
DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 60

QY 61 LCGSHLVKALYVCGERGFFYTPKTRGIVEQCCSTICSLYLENYCN 107
DB 61 LCGSHLVKALYVCGERGFFYTPKTRGIVEQCCSTICSLYLENYCN 107

RESULT 2
US-10-054-873-7

Sequence 7; Application US/10054873
Publication No. US20020164712A1

GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054, 873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000

ATTORNEY/AGENT INFORMATION:
NAME: MYCROFT, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-054-873-7

Query Match 94.6%; Score 555.5; DB 13; Length 150;
Best Local Similarity 71.3%; Pred. No. 5, 7e-57;
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 49
DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 60

QY 50 -----LGTGRFVNOHLCGSHLVKALYVCGER 77
DB 61 TBSNREFTQGXNLELRLSILLIGSWLEPVLQGTGRFVNOHLCGSHLVKALYVCGER 120
QY 78 GPEYTPKTRGIVEQCCSTICSLYLENYCN 107
DB 121 GPEYTPKTRGIVEQCCSTICSLYLENYCN 150

RESULT 3
US-10-101-454-39
Sequence 39; Application US/10101454
Publication No. US20040110664A1

GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Id
Andersen, Asger Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101, 454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400, 256
FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985, 220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 137 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-10-101-454-39

Query Match 51.5%; Score 302.5; DB 16; Length 137;
Best Local Similarity 50.0%; Pred. No. 2, 4e-27;
Matches 70; Conservative 4; Mismatches 27; Indels 39; Gaps 4;

QY 2 FPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 48
DB 3 FPTPLSRFLPDNMLRAHRLHQLAFDTYQFEFEAYIPKQKXSFLONPFGSPFVNH 57
QY 49 PLGTG-----PFTVNOHLCGSHLVKALYVCGERGFFYTPKTRG 87
DB 58 STNGLLFTNTTIAIAKEGVSMKRFVNHLCGSHLVKALYVCGERGFFYTPKTRG 117
QY 88 IVEQCCSTICSLYLENYCN 107
DB 118 IVEQCCSTICSLYLENYCN 137

RESULT 4
US-10-101-454-45
Sequence 45, Application US/10101454
Publication No. US20040110664A1
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Ib
Andersen, Aaser Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSER: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 145 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-101-454-45
Query Match 50.9%; Score 299; DB 16; Length 145;
Best Local Similarity 100.0%; Pred. No. 6,7e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 RFVNHLCGSHLVEALYLVCGERGFFYPTKRTGIVEQCCTSIICSLYOLENYCN 107
DB 93 RFVNHLCGSHLVEALYLVCGERGFFYPTKRTGIVEQCCTSIICSLYOLENYCN 145
RESULT 5
US-10-101-454-48
Sequence 48, Application US/10101454
Publication No. US20040110664A1
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, Ib
Andersen, Aaser Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSER: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York

COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985,220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-101-454-48
Query Match 50.9%; Score 299; DB 16; Length 146;
Best Local Similarity 100.0%; Pred. No. 6,7e-27;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 RFVNHLCGSHLVEALYLVCGERGFFYPTKRTGIVEQCCTSIICSLYOLENYCN 107
DB 94 RFVNHLCGSHLVEALYLVCGERGFFYPTKRTGIVEQCCTSIICSLYOLENYCN 146
RESULT 6
US-10-054-873-5
Sequence 5, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-10-054-873-5
Query Match 50.1%; Score 294; DB 13; Length 52;
Best Local Similarity 100.0%; Pred. No. 7.4e-27;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 PFVNHGCGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGSLYLENYCN 107
DB 1 PFVNHGCGSHLVEALYLVCGERGFFYTPKTRGIVECCCTSGSLYLENYCN 52

RESULT 7
US-09-894-711-18
Sequence 18, Application US/09894711
Patent No. US2002013714A1
GENERAL INFORMATION:
APPLICANT: Kjeldsen, Thomas Borglum
TITLE OF INVENTION: Method for making insulin precursors and
TITLE OF INVENTION: insulin precursor analogues having improved fermentation
FILE REFERENCE: 6148.400-US
CURRENT APPLICATION NUMBER: US/09/894, 711
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: PA 2000 00443
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: PA 1999 01869
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 60/211,081
PRIOR FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: 60/181,450
PRIOR FILING DATE: 2000-02-10
PRIOR APPLICATION NUMBER: 09/740,359
PRIOR FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 18
LENGTH: 124
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-894-711-18

Query Match 48.5%; Score 284.5; DB 9; Length 124;
Best Local Similarity 92.7%; Pred. No. 2.8e-25;
Matches 51; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 54 PFVNHGCGSHLVEALYLVCGERGFFYTPK-TRGIVECCCTSGSLYLENYCN 107
DB 70 PFVNHGCGSHLVEALYLVCGERGFFYTPKAKGIVECCCTSGSLYLENYCN 124

RESULT 8
US-09-861-687-19
Sequence 19, Application US/09861687
Publication No. US20020193292A1
GENERAL INFORMATION:
APPLICANT: Markussen, Jan
Jonassen, Ib
Havelund, Svend
Brandt, Jakob
Kurtzhals, Peter
Hansen, Hertz Per
Kaarsholm, Niels Christian

TITLE OF INVENTION: INSULIN DERIVATIVES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: No. US20020193292A10 No. US20020193292A1disk of No. US200201932

STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/861,687
FILING DATE: 21-May-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/932,082
FILING DATE: 16-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4341,204-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-861-687-19
Query Match 48.4%; Score 284; DB 9; Length 138;
Best Local Similarity 48.2%; Pred. No. 3.6e-25;
Matches 68; Conservative 5; Mismatches 28; Indels 40; Gaps 5;

QY 2 PFVNHGCGSHLVEALYLVCGERGFFYTPK-TRGIVECCCTSGSLYLENYCN 107
DB 3 PFVNHGCGSHLVEALYLVCGERGFFYTPK-TRGIVECCCTSGSLYLENYCN 107
QY 49 PLGTG-----PFVNHGCGSHLVEALYLVCGERGFFYTPK-TR 86
DB 56 STNGHLFINTTASIAKEGVSLDRPFVNHGCGSHLVEALYLVCGERGFFYTPKAK 117
QY 87 GIVECCCTSGSLYLENYCN 107
DB 118 GIVECCCTSGSLYLENYCN 138

RESULT 9
US-10-620-651-19
Sequence 19, Application US/10620651
Publication No. US20040067874A1
GENERAL INFORMATION:
APPLICANT: Markussen, Jan
Jonassen, Ib
Havelund, Svend
Brandt, Jakob
Kurtzhals, Peter
Hansen, Hertz Per
Kaarsholm, Niels Christian
TITLE OF INVENTION: INSULIN DERIVATIVES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: No. US20040067874A10 No. US20040067874A1disk of No. US2004006
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York

Sat Nov 6 18:59:29 2004

us-10-054-873-6.rapb

Page 5

COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/620,651
FILING DATE: 15-Jul-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/932,082
FILING DATE: 17-SEPT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4341.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-620-651-19
Query Match 48.4%; Score 284; DB 15; Length 138;
Best Local Similarity 48.2%; Pred. No. 3.6e-25;
Matches 68; Conservative 5; Mismatches 28; Indels 40; Gaps 5;
QY 2 PPTPLSLRPLDNLRAHLQLAFTYQEFEEAVIPKEQ--KYSFLO-----N 48
DB 3 FPSI-----FTAVLPAASSALAPVNTTDEDTAQIPAAVIGYSDLDGDFVAVLPFSN 57
QY 49 PLGTG-----PRFVNOHLCGSHLYEALYLVGGERGFYTPKT--TR 96
DB 58 STNNGLFPIINTTIASIAKEBGVSLDKRFVNOHLCGSHLYEALYLVGGERGFYTPKAK 117
QY 87 GIVEOCCSTICSLYOLENYCN 107
DB 118 GIVEOCCSTICSLYOLENYCN 138
RESULT 10
US-10-101-454-33
Sequence 33, Application US/10101454
Publication No. US20040110664A1
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, ID
Andersen, Asger Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454

FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 140 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-10-101-454-33
Query Match 48.4%; Score 284; DB 16; Length 140;
Best Local Similarity 47.6%; Pred. No. 3.7e-25;
Matches 68; Conservative 6; Mismatches 27; Indels 42; Gaps 5;
QY 2 PPTPLSLRPLDNLRAHLQLAFTYQEFEEAVIPKEQ--KYSFLO-----N 48
DB 3 FPSI-----FTAVLPAASSALAPVNTTDEDTAQIPAAVIGYSDLDGDFVAVLPFSN 57
QY 49 PLGTG-----PRFVNOHLCGSHLYEALYLVGGERGFYTPKT-- 85
DB 58 STNNGLFPIINTTIASIAKEBGVSLDKRFVNOHLCGSHLYEALYLVGGERGFYTPKSD 117
QY 86 -RGIVEOCCSTICSLYOLENYCN 107
DB 118 AKGIVEOCCSTICSLYOLENYCN 140
RESULT 11
US-10-101-454-42
Sequence 42, Application US/10101454
Publication No. US20040110664A1
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
Halstrom, John
Jonassen, ID
Andersen, Asger Sloth
Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/101,454
FILING DATE: 20-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/400,256
FILING DATE: 03-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 140 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 42:

US-10-101-454-42

Query Match

48.4%; Score 284; DB 16; Length 140;

Best Local Similarity

47.6%; Pred. No. 3.7e-25;

Matches

68; Conservative 6; Mismatches 27; Indels 42; Gaps 5;

QY

2 FFTPLSLPDMALRLHQLAFDYQSEEAATPKKQ--KVSFLQ-----N 48

DB

3 FFSI-----FPAVFAASSALAPAVNTTDETRQIPAEAVIGYSLSDPFAVALPFSN 57

QY

49 PLGTG-----PRVNOHLGSHLYEALYVCGERGFFYTPKT-- 85

DB

58 STNNGLPFINTTASIAKEGVSMARFVNOHLGSHLYEALYVCGERGFFYTPKSD 117

QY

86 -RGVVEOCCSTICSLYLENYCN 107

DB

118 AKGIVEOCCSTICSLYLENYCN 140

RESULT 12

US-10-101-454-15

Sequence 15, Application US/10101454

Publication No. US20040110664A1

GENERAL INFORMATION:

APPLICANT: Havelund, Svend

Halstrom, John

Jonassen, Ib

Jonassen, Asger Sloth

Markussen, Jan

TITLE OF INVENTION: ACYLATED INSULIN

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESSES:

ADDRESS: Novo Nordisk of North America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/101/101,454

FILING DATE: 20-Mar-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/400,256

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3985,220-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 104 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-10-101-454-15

Query Match

47.9%; Score 281; DB 16; Length 104;

Best Local Similarity

71.8%; Pred. No. 5.8e-25;

Matches

56; Conservative 6; Mismatches 8; Indels 8; Gaps 3;

QY

37 IPKQ---KVSFLONPLGTGPRFVNOHLGSHLYEALYVCGERGFFYTPKT--RGIV 89

DB

28 IPESLIIAENTTLAN-VAMAKRFVNOHLGSHLYEALYVCGERGFFYTPKSDAKGIV 86

QY

90 EQCCTSTICSLYLENYCN 107

DB

87 EQCCTSTICSLYLENYCN 104

RESULT 13

US-09-858-935B-5

Sequence 5, Application US/09858935B

Publication No. US20030069177A1

GENERAL INFORMATION:

APPLICANT: Dubaigle, Yves

APPLICANT: Filvaroff, Ellen

APPLICANT: Lomman, Henry B.

FILE REFERENCE: P1794R1

CURRENT APPLICATION NUMBER: US/09/858,935B

CURRENT FILING DATE: 2002-07-02

PRIOR APPLICATION NUMBER: US 60/248,985

PRIOR FILING DATE: 2000-11-15

PRIOR APPLICATION NUMBER: US 60/204,490

PRIOR FILING DATE: 2000-05-16

NUMBER OF SEQ ID NOS: 153

SEQ ID NO 5

LENGTH: 51

TYPE: PRT

ORGANISM: Homo sapiens

US-09-858-935B-5

Query Match

47.4%; Score 278.5; DB 10; Length 51;

Best Local Similarity

98.1%; Pred. No. 4.8e-25;

Matches

51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY

56 FVNOHLGSHLYEALYVCGERGFFYTPKTGIVEOCCSTICSLYLENYCN 107

DB

1 FVNOHLGSHLYEALYVCGERGFFYTPKT-GIVEOCCSTICSLYLENYCN 51

RESULT 14

US-10-028-410-3

Sequence 3, Application US/10028410

Publication No. US20020160955A1

GENERAL INFORMATION:

APPLICANT: Dubaigle, Yves

APPLICANT: Lomman, Henry

TITLE OF INVENTION: PROTEIN VARIANTS

FILE REFERENCE: P1712R1-1

CURRENT APPLICATION NUMBER: US/10/028,410

CURRENT FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: US/09/477,924

PRIOR FILING DATE: 2000-01-05

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 3

LENGTH: 51

TYPE: PRT

ORGANISM: Homo sapiens

US-10-028-410-3

Query Match

47.4%; Score 278.5; DB 13; Length 51;

Best Local Similarity

96.1%; Pred. No. 4.8e-25;

Matches

51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY

56 FVNOHLGSHLYEALYVCGERGFFYTPKTGIVEOCCSTICSLYLENYCN 107

Sat Nov 6 18:59:29 2004

us-10-054-873-6.rapb

Page 7

Db 1 FVNHLCGSHVEALYLVCGERGFFYTPKT-GIVEOCCTISCSLYOLENYCN 51

RESULT 15
US-10-444-326-3
; Sequence 3, Application US/10444326
; Publication No. US20030191065A1
; GENERAL INFORMATION:
; APPLICANT: Dubague, Yves
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1
; CURRENT APPLICATION NUMBER: US/10/444,326
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US/09/723,866
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US/09/477,923
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-444-326-3

Query Match 47.4%; Score 278.5; DB 14; Length 51;
Best Local Similarity 98.1%; Pred. No. 4.8e-25;
Matches 51; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 56 FVNHLCGSHVEALYLVCGERGFFYTPKT-GIVEOCCTISCSLYOLENYCN 107
Db 1 FVNHLCGSHVEALYLVCGERGFFYTPKT-GIVEOCCTISCSLYOLENYCN 51

Search completed: November 2, 2004, 20:59:23
Job time : 81.1513 secs

This Page Blank (uspto)

Sat Nov 6 18:59:31 2004

us-10-054-873-6.rup

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:48:36 ; Search time 103.644 Seconds
(without alignments)
594.006 Million cell updates/sec

Title: US-10-054-873-6

Perfect score: 587

Sequence: 1 MFPTPLSRLEFDNMLRAHR.....IVGQCTSTICSLYQLENYCN 107

Scoring table:

BLASTSUM62
Gapop 10.0, Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: uniprot_sprot.*
2: uniprot_trmb1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	275	46.8	96	Q7M0U6
2	273.5	46.6	51	INS_BALPH
3	273.5	46.6	51	INS_ELEMA
4	273	46.5	110	INS_CERA
5	273	46.5	110	INS_MACRA
6	271.5	46.3	51	Q7M0G1
7	268.5	45.7	51	INS_ACOC
8	267.5	45.6	51	Q7M217
9	267	45.5	110	INS_HUMAN
10	267	45.5	110	INS_PANTR
11	267	45.5	110	INS_PONRY
12	267	45.5	110	O6YK33
13	267	45.5	110	AA035454
14	267	45.5	110	AA006935
15	267	45.5	110	AA039451
16	266	45.3	110	INS_SPER
17	263.5	44.9	51	INS_BALBO
18	263.5	44.9	51	INS_CAMBR
19	263.5	44.9	51	INS_CAMPH
20	263	44.8	108	INS_PIG
21	263	44.8	108	AA000952
22	263	44.8	108	AA000954
23	263	44.8	108	AA000957
24	263	44.8	108	AA000960
25	263	44.8	108	AA000963
26	263	44.8	108	AA000966
27	263	44.8	108	AA000969
28	263	44.8	108	AA000972
29	263	44.8	108	AA000975
30	263	44.8	108	AA000978
31	263	44.8	108	AA000981

32	263	44.8	108	AA000983	AA000983 sus scrof
33	263	44.8	108	AA000985	AA000985 sus scrof
34	263	44.8	108	AA000987	AA000987 sus scrof
35	263	44.8	108	AA000990	AA000990 sus scrof
36	263	44.8	110	INS_RABIT	P01311 Oryctolagus
37	262.5	44.7	51	INS_FELCA	P06306 felis silve
38	262	44.6	110	INS_CANFA	P01321 canis fam1
39	260	44.3	105	INS_BOVIN	P01317 bos taurus
40	258.5	44.0	105	INS_AOTTR	P10604 actus trivl
41	257	43.8	110	INS_PSAOB	O62587 psammomyx o
42	257	43.8	51	INS_DIMA	P18109 didelphis m
43	256.5	43.7	217	INS_HUMAN	P01241 homo sapien
44	255.5	43.5	217	SOMA_MACMU	P33093 macaca mula
45	255.5	43.5	217	SOMA_MACMU	P33093 macaca mula

ALIGNMENTS

RESULT 1

ID	Q7M0U6	PRELIMINARY;	PRT;	96 AA.
AC	Q7M0U6			
DT	01-MAR-2004 (TREMURel. 26, Created)			
DT	01-MAR-2004 (TREMURel. 26, Last sequence update)			
DT	01-MAR-2004 (TREMURel. 26, Last annotation update)			
DE	Epidermal growth factor/single chain insulin fusion protein (Fragment)			
OS	Bacillus brevis (Brevibacillus brevis)			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Brevibacillus			
OX	NCBI_TaxID=1393;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEHLIN=20335834; PubMed=10879487;			
RA	Koh M., Hanagata H., Ebisu S., Morihara K., Takagi H.;			
RT	"Use of Bacillus brevis for synthesis and secretion of Des-B30 single-chain human insulin precursor".			
RL	BioSci. Biotechnol. Biochem. 64:1079-1081(2000).			
DR	PIR; PC7082; PC7082.			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0005179; F:hormone activity; IEA.			
DR	GO; GO:0007582; P:physiological process; IEA.			
DR	InterPro; IPRO04825; Ins/IGF/relax.			
DR	Pfam; PF00049; Insulin; 1.			
DR	PRINTS; PR00277; INSULIN.			
DR	PROSITE; PS00262; INSULIN; 1.			
FT	NON_TER			
FT	NON_TER			
SO	SEQUENCE			
Query Match	96 AA; 10473 MW; 4505D710C289092A CRC64;			
Best Local Similarity	94.3%; Score 275; DB 2; Length 96;			
Matches	50; Conservative 1; Mismatches 0; Indels 2; Gaps 1;			
Cy	55 RPYNHGSHLVEALYLVCGERGFFYPTKRTGIYDQCTSLCSLYQLENYCN 107			
Db	46 KPVNHGSHLVEALYLVCGERGFFYPTK--GIVEQCTSTICSLYQLENYCN 96			
RESULT 2				
ID	INS_BALPH	STANDARD;	PRT;	51 AA.
AC	P01312			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Insulin.			
GN	Name=INS;			
OS	Balaenoptera physalus (Finback whale) (Common rorqual), and			
OS	Physeter carodon (Sperm whale) (Physeter macrocephalus).			
OC	Balaenoptera; Cetartiodactyla; Cetacea; Mysticeti;			
OC	Balaenoptera; Cetartiodactyla; Cetacea; Mysticeti;			
OC	Balaenoptera; Cetartiodactyla; Cetacea; Mysticeti;			

```

OK NCB1_Taxid=9770, 9755;
RN [1]
RN PARTIAL SEQUENCE.
RC SPECIES=B. physalus;
RA Hama H., Titani K., Sakaki S., Narita K.;
RT "The amino acid sequence in fin-whale insulin.";
RN J. Biochem. 56:285-293(1964).
RN [2]
RN SEQUENCE.
RC SPECIES=P. carodon;
RA Ishihara Y., Saito T., Ito Y., Fujino M.;
RT "Structure of sperm- and sei-whale insulins and their breakdown by
  whale pepsin.";
RN Nature 181:1468-1469(1958).
RN [3]
RN SEQUENCE.
RC SPECIES=P. carodon;
RA Harris J., Sanger F., Naughton M.A.;
RT "Species differences in insulin.";
RN Arch. Biochem. Biophys. 65:427-438(1956).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
  increases cell permeability to monosaccharides, amino acids and
  fatty acids. It accelerates glycolysis, the pentose phosphate
  cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
  disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
DR PIR: A91918; INWHF.
DR PIR: A91918; INWHF.
DR HSSP: P01317; IAHF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR PRINTS: PR00277; INSULINB.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
DR Direct protein sequencing; Glucose metabolism; Hormone;
KM Insulin family.
FT CHAIN 1 30 Insulin B chain.
FT NON_CONS 30 31
FT CHAIN 31 51 Insulin A chain.
FT DISULFID 7 37 Interchain.
FT DISULFID 19 50 Interchain.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5766 MW; 9007B514691A7CDD CRC64;

Query Match 46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 96.2%; Pred. No. 3.6e-21;
Matches 50; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYLVCGERGFFYTPKRTGIVECCCTGSLVQLNYYNCN 107
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKRTGIVECCCTGSLVQLNYYNCN 51

RESULT 3
INS_ELEMNA STANDARD; PRT; 51 AA.
AC P01316;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin.
OS Name=INS;
OS Elephas maximus (Indian elephant).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Proboscidea; Elephantidae; Elephae.
OX NCB1_Taxid=9783;
RN [1]
RN SEQUENCE.
RA MEDLINE=66160119; Pubmed=5945593;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RN Am. J. Med. 40:662-666(1966).

```

```

CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
  increases cell permeability to monosaccharides, amino acids and
  fatty acids. It accelerates glycolysis, the pentose phosphate
  cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
  disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: The species of elephant is not given, but it is
  most probably the Indian elephant (Elephas maximus).
CC -1- SIMILARITY: Belongs to the insulin family.
DR HSSP: P01308; IAHF.
DR InterPro: IPR004825; Ins/IGF/relax.
DR PRINTS: PR00277; INSULINB.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
DR Direct protein sequencing; Glucose metabolism; Hormone;
KM Insulin family.
FT CHAIN 1 30 Insulin B chain.
FT NON_CONS 30 31
FT CHAIN 31 51 Insulin A chain.
FT DISULFID 7 37 Interchain.
FT DISULFID 19 50 Interchain.
FT DISULFID 36 41
SQ SEQUENCE 51 AA; 5752 MW; 9007B50CDB457D6D CRC64;

Query Match 46.6%; Score 273.5; DB 1; Length 51;
Best Local Similarity 94.2%; Pred. No. 3.6e-21;
Matches 49; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 56 FVNHLCGSHLVEALYLVCGERGFFYTPKRTGIVECCCTGSLVQLNYYNCN 107
DB 1 FVNHLCGSHLVEALYLVCGERGFFYTPKRTGIVECCCTGSLVQLNYYNCN 51

RESULT 4
INS_ELEMNA STANDARD; PRT; 110 AA.
AC P30407; P01309;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin precursor.
OS Name=INS;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCB1_Taxid=9534;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=92219953; Pubmed=1560757;
RA Saito S., Bell G.I., Li W.;
RT "Sequences of primate insulin genes support the hypothesis of a slower
  rate of molecular evolution in humans and apes than in monkeys.";
RN Mol. Biol. Evol. 9:193-203(1992).
RN [2]
RN SEQUENCE OF 57-87.
RA MEDLINE=72258016; Pubmed=4626369;
RA Peterson J.D., Nehrllich S., Oyer P.B., Steiner D.F.;
RT "Determination of the amino acid sequence of the monkey, sheep, and
  dog proinsulin C-peptides by a semi-micro Edman degradation
  procedure.";
RN J. Biol. Chem. 247:4865-4871(1972).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
  increases cell permeability to monosaccharides, amino acids and
  fatty acids. It accelerates glycolysis, the pentose phosphate
  cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
  disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X61092; CAA43405.1; .
CC PIR: B42179; B42179.
CC HSSP: P01308; IATD.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Direct protein sequencing; Glucose metabolism; Hormone;
KM Insulin family; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 Insulin B chain.
FT PROPEP 57 87 C peptide.
FT CHAIN 90 110 Insulin A chain.
FT DISULFID 31 96 Interchain.
FT DISULFID 43 109 Interchain.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12019 MW; 95A1F54B7E247F9 CRC64;

Query Match 46.5%; Score 273; DB 1; Length 110;
Best Local Similarity 60.2%; Pred. No. 9.1e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PFVNQHLGSHLVYALVYCGERGFFYTPK----- 85
DB 23 PAFVNQHLGSHLVYALVYCGERGFFYTPKREADPQVGVELGGPGAGSLQPLAL 82
QY 86 -----RGIVECCCTSGISLYOLENYCN 107
DB 83 EGSLOKRGIVECCCTSGISLYOLENYCN 110

RESULT 5
INS_MACPA STANDARD; PRT; 110 AA.
ID INS_MACPA
AC P30406; P01309; .
DT 21-JUL-1986 (Rel. 01. Created).
DT 13-AUG-1987 (Rel. 05; Last sequence update)
DT 05-JUL-2004 (Rel. 44; Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=83080474; PubMed=6184262;
RA Wetekamp M., Gronenberg J., Leineweber M., Wengenmayer F.,
RA Winnacker E.-L.;
RT "The nucleotide sequence of cDNA coding for preproinsulin from the
RT primate Macaca fascicularis."
RT Gene 19:179-183(1982).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycogenesis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00336; AAA36849.1; .
CC PIR: J00178; J00178.
CC HSSP: P01308; IATD.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR SMART: SM00078; IIGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Glucose metabolism; Hormone; Insulin family; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 Insulin B chain.
FT PROPEP 57 87 C peptide.
FT CHAIN 90 110 Insulin A chain.
FT DISULFID 31 96 Interchain.
FT DISULFID 43 109 Interchain.
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 11991 MW; 83C6E33A80A420F9 CRC64;

Query Match 46.5%; Score 273; DB 1; Length 110;
Best Local Similarity 60.2%; Pred. No. 9.1e-21;
Matches 53; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 54 PFVNQHLGSHLVYALVYCGERGFFYTPK----- 85
DB 23 PAFVNQHLGSHLVYALVYCGERGFFYTPKREADPQVGVELGGPGAGSLQPLAL 82
QY 86 -----RGIVECCCTSGISLYOLENYCN 107
DB 83 EGSLOKRGIVECCCTSGISLYOLENYCN 110

RESULT 6
Q7MOG1 PRELIMINARY; PRT; 51 AA.
ID Q7MOG1
AC Q7MOG1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Insulin
OS Cricetidae sp. (Hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae.
OX NCBI_TaxID=36483;
RN (1)
RP SEQUENCE.
RA Neelson F.A., Deicher H.K., Steinman H., Lebovitz H.E.;
RT "Structure of hamster insulin: comparison with a tumor insulin."
RT Fed. Proc. 32:360-360(1973).
CC -1- SIMILARITY: Belongs to the insulin family.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -----
CC PIR: A91456; A91456.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0005179; F:hormone activity; IEA.
DR GO: GO:0007582; P:physiological process; IEA.
DR InterPro: IPR004825; Ins/IGF/relax.
DR Pfam: PF00049; Insulin; 1.
DR PRINTS: PR00277; INSULIN.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family.
SQ SEQUENCE 51 AA; 5768 MW; 9006B6469047D3D CRC64;

Query Match 46.3%; Score 271.5; DB 2; Length 51;
Best Local Similarity 94.2%; Pred. No. 5.9e-21;
Matches 49; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 56 PFVNQHLGSHLVYALVYCGERGFFYTPKTRIVECCCTSGISLYOLENYCN 107
DB 1 PFVNQHLGSHLVYALVYCGERGFFYTPKRS-GIVQCCCTSGISLYOLENYCN 51

RESULT 7
INS_ACOCA STANDARD; PRT; 51 AA.
AC P01324;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Insulin.
GN Name=INS;
OS Acomys caliginus (Egyptian spiny mouse).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Acomys.
OX NCBI_Taxid=10068;
RN [1]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=72189454; PubMed=5028210;
RA Buzenli H.F., Humbel R.E.,
RT "Isolation and partial structural analysis of insulin from mouse (Mus musculus) and spiny mouse (Acomys caliginus)."
RT Hoppe-Seyler's Z. Physiol. Chem. 353:444-450(1972).
CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.
CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two disulfide bonds.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the insulin family.
DR HSPB; P01308; 1E6.
DR InterPro; IPR004825; Ins/IGF/relax.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IIGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KM Direct protein sequencing; Glucose metabolism; Hormone.
KM Insulin family.
FT CHAIN 1 30 Insulin B chain.
FT NON CONS 30 31
FT CHAIN 31 51 Insulin A chain.
FT DISULFID 7 37 Interchain (By similarity).
FT DISULFID 19 50 Interchain (By similarity).
FT DISULFID 36 41 By similarity.
SQ SEQUENCE 51 AA; 5768 MW; 992BDB629047D3D CRC64;
Query Match 45.7%; Score 268.5; DB 1; Length 51;
Best Local Similarity 92.3%; Pred. No. 1.2e-20;
Matches 48; Conservative 3; Mismatches 0; Indels 1; Gaps 1;
QY 56 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVECCCTGSLYLENYCN 107
DB 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKS-GIVDCCCTGSLYLENYCN 51
RESULT 8
QY 7M217 PRELIMINARY; PRT; 51 AA.
AC Q7M217;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Insulin precursor (Fragments).
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Euphorbia; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_Taxid=3823;
RN [1]
RP SEQUENCE.
RA Oliveira A.E.A., Machado O.L.T., Gomes V.M., Xavier-Neto J.,
RA Pereira A.C.P., Vieira J.G.H., Fernandes K.V.S., Xavier-Filho J.,
RN "Jack bean seed coat contains a protein with complete sequence

RT homology to bovine insulin.";
RL Protein Pept. Lett. 6:15-21(1999).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Belongs to the insulin family.
DR PIR; B59151; B59151.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR GO; GO:0007582; P:physiological process; IEA.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR PROSITE; PS00262; INSULIN; 1.
KM Insulin family.
FT CHAIN 1 51
FT NON TER 1 1
FT NON TER 51 51
SQ SEQUENCE 51 AA; 5722 MW; 9007B50CCA0A7DD CRC64;
Query Match 45.6%; Score 267.5; DB 2; Length 51;
Best Local Similarity 92.3%; Pred. No. 1.5e-20;
Matches 48; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
QY 56 FVNOHLCGSHLVEALYLVCGERGFFYTPKTRGIVECCCTGSLYLENYCN 107
DB 1 FVNOHLCGSHLVEALYLVCGERGFFYTPKA-GIVECCASVCSLYLENYCN 51
RESULT 9
INS_HUMAN STANDARD; PRT; 110 AA.
AC P01308;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Insulin precursor.
GN Name=INS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80120725; PubMed=6243748;
RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
RA Goodman H.M.;
RT "Sequence of the human insulin gene.";
RL Nature 284:26-32(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=80236313; PubMed=6248962;
RA Ullrich A., Dull T.J., Gray A., Brosius J., Sures I.;
RT "Genetic variation in the human insulin gene.";
RL Science 209:612-615(1980).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80054779; PubMed=503234;
RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
RA Rutter W.J.;
RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin.";
RL Nature 282:525-527(1979).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=80147417; PubMed=6927840;
RA Sures I., Goeddel D.V., Gray A., Ullrich A.;
RT "Nucleotide sequence of human preproinsulin complementary DNA.";
RL Science 208:57-59(1980).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=93364428; PubMed=8358440;
RA Lucassen A.M., Bell J., Jullier C., Lathrop M.;
RT "Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR.";
RL Nat. Genet. 4:305-310(1993).
RN [6]

RP SEQUENCE FROM N.A.
RX TISSUE=pancreas;
RX MEDLINE=22388577; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.H., Felingold E.A., Grouse L.H., Dexe J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stempson M., Soares M.B., Bonaldo M.F., Casavant T.L., Scherz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Raba S., Longuelo N.A., Peters G.J., Abramson R.D., Mulady S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maier M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE OF 1-59 FROM N.A.
RC TISSUE=Blood;
RA Fajardy I.T., Weill J.J., Stuckens C.C., Danze P.M.P.,
RT "Description of a novel RRP diallelic polymorphism (-127 Bsgt C/G)
RT within the 5' region of insulin gene";
RL submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
RP SEQUENCE OF 25-54 AND 90-110.
RA Nicol D.S.H.W., Smith L.F.,
RT "Amino-acid sequence of human insulin";
RL Nature 187:483-485 (1960).
RP SEQUENCE OF 57-87.
RX MEDLINE=7116410; PubMed=5101771;
RA Oyer P.E., Cho S., Peterson J.D., Steiner D.F.,
RT "Studies on human proinsulin. Isolation and amino acid sequence of the
RT human pancreatic C-peptide.";
RL J. Biol. Chem. 246:1375-1386(1971).
RP SEQUENCE OF 57-87.
RX MEDLINE=7125722; PubMed=5560404;
RA KO A., Smyth D.G., Markussen J., Sundby F.,
RT "The amino acid sequence of the C-peptide of human proinsulin.";
RL Eur. J. Biochem. 20:190-199(1971).
RP SYNTHESIS.
RX MEDLINE=7507277; PubMed=4443293;
RA Sieber P., Kanher B., Hartmann A., Joehl A., Rinkler B., Rittel W.,
RT "Total synthesis of human insulin under directed formation of the
RT disulfide bonds.";
RL Helv. Chim. Acta 57:2617-2621(1974).
RP SYNTHESIS OF 57-87.
RX MEDLINE=7504007; PubMed=4803504;
RA Naichani V.K.,
RT "Studies on polypeptides, IV. The synthesis of C-peptide of human
RT proinsulin.";
RL Hoppe-Seyler's Z. Physiol. Chem. 354:655-672(1973).
RP SYNTHESIS OF 65-69 AND 70-73.
RX MEDLINE=73161263; PubMed=4698555;
RA Geiger R., Volk A.,
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). 3. Synthesis of the sequences 14-17 and 9-13 of
RL human proinsulin C peptides.";
RL Chem. Ber. 106:199-205(1973).
RP SYNTHESIS OF 84-87.
RX MEDLINE=73161261; PubMed=4698553;
RA Geiger R., Jaeger G., Keonig W., Trenth G.,
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hc peptide). 1. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide.";
RL Chem. Ber. 106:188-192(1973).
RP VARIANT LOS ANGELES SER-48.
RX MEDLINE=84010233; PubMed=6312455;
RA Harenda M., Chan S.U., Kwok S.C.M., Rubenstein A.H., Steiner D.F.,
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB34]insulin.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
RP VARIANT LOS ANGELES SER-48 AND CHICAGO LEU-49.
RX MEDLINE=84170233; PubMed=6424111;
RA Shoelson S., Flickova M., Harenda M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.,
RT "Identification of a mutant human insulin predicted to contain a
RT serine-for-phenylalanine substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RP VARIANT PROVIDENCE ASP-34.
RX MEDLINE=87175640; PubMed=3470784;
RA Chan S.U., Seino S., Gruppiso P.A., Schwartz R., Steiner D.F.,
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2197(1987).
RP VARIANT MAKAYAMA LEU-92.
RX MEDLINE=87056122; PubMed=3537011;
RA Sakura H., Iwamoto Y., Sakamoto Y., Kizuya T., Hirata H.,
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A3 (Val-->Leu) isolated from the pancreas.";
RL J. Clin. Invest. 78:1666-1672(1986).
RP VARIANT HIS-89.
RX MEDLINE=80317021; PubMed=2196279;
RA Barbetti F., Raden N., Kadowaki T., Cama A., Accili D., Gabbay K.H.,
RA Wernech J.A., Taylor S.T., Koch J.,
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RL J. Clin. Endocrinol. Metab. 71:164-169(1990).
RP VARIANT HIS-89.
RX MEDLINE=85261996; PubMed=4019786;
RA Shibaaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.,
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RL J. Clin. Invest. 76:378-380(1985).
RP VARIANT KYOTO LEU-89.
RX MEDLINE=92291307; PubMed=1601997;
RA Iano H., Kitano N., Morimoto M., Polonsky K.S., Imura H., Seino Y.,
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto).";
RL J. Clin. Invest. 89:1902-1907(1992).
RP STRUCTURE BY NMR.
RX MEDLINE=9104966; PubMed=2271664;
RA Hua O.-X., Weiss M.A.,
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide analogue and comparison with crystal
RL Biochemistry 29:10545-10555(1990).
RP STRUCTURE BY NMR.
RX MEDLINE=91242467; PubMed=2036420;
RA Hua O.-X., Weiss M.A.,
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein

RT dynamics and receptor recognition.",
 RL Biochemistry 30:5505-5515(1991).
 RN [24]
 RP STRUCTURE BY NMR.
 RX MEDLINE=91265527; PubMed=164635;
 RA Hua Q.-X., Weiss M.A.;
 RT "Two-dimensional NMR studies of Des-(B26-B30)-insulin: sequence-specific resonance assignments and effects of solvent composition.",
 RL Biochim. Biophys. Acta 1078:101-110(1991).

Query Match 45.5%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDLQVGYVELGGGPGAGSLQPLALEG 84
 QY 86 ---RGIVECCCTSCISLYOLENYCN 107
 DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 10
 INS_PANTR STANDARD; PRT; 110 AA.
 AC P30410;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92219953; PubMed=1560757;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; X61089; CAA43403.1; -
 DR EMBL; AY137497; AAN06933.1; -
 DR PIR; A42179; A42179.
 DR HSSP; P01308; IAI0.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin.1.
 DR PRINTS; PR00277; INSULINB.

DR PROSITE; PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Insulin A chain.
 FT DISULFID 43 109 Interchain.
 FT DISULFID 95 100 Interchain.
 FT HELIX 32 43
 FT TURN 60 61
 FT TURN 60 71
 FT HELIX 65 71
 FT HELIX 82
 SQ SEQUENCE 110 AA; 12025 MW; 41EB8DF79837CEFS CRC64;

Query Match 45.5%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNHLCGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNHLCGSHLVEALYLVCGERGFFYTPKTRRAEDLQVGYVELGGGPGAGSLQPLALEG 84
 QY 86 ---RGIVECCCTSCISLYOLENYCN 107
 DB 85 SLQKRGIVECCCTSCISLYOLENYCN 110

RESULT 11
 INS_PONPY STANDARD; PRT; 110 AA.
 AC Q8HXV2;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Insulin precursor.
 GN Name=INS;
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878; DOI=10.1101/gr.948003;
 RA Stead J.D.H., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 CC -1- FUNCTION: Insulin decreases blood glucose concentration. It increases cell permeability to monosaccharides, amino acids and fatty acids. It accelerates glycolysis, the pentose phosphate cycle, and glycogen synthesis in liver.
 CC -1- SUBUNIT: Heterodimer of a B chain and an A chain linked by two disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to the insulin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; AY137503; AAN06937.1; -
 DR HSSP; P01308; IAI0.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin.1.
 DR PRINTS; PR00277; INSULINB.
 DR SMART; SMO0078; IIGF.1.
 DR PROSITE; PS00262; INSULIN; 1.
 KW Glucose metabolism; Hormone; Insulin family; Signal.
 FT SIGNAL 1 24 By similarity.

FT CHAIN 25 54 Insulin B chain.
 FT PROPEP 57 87 C-peptide.
 FT CHAIN 90 110 Insulin A chain.
 FT DISULFID 31 96 Interchain (By similarity).
 FT DISULFID 43 109 Interchain (By similarity).
 FT DISULFID 95 100 By similarity.
 SQ SEQUENCE 110 AA; 12038 MW; 22D2B32B94F520F8 CRC64;

Query Match 45.5%; Score 267; DB 1; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREARDLVQGVELGGPGAGSLQPLALEG 84
 QY 86 ----RGIVEOCCTSIICSLYOLENYCN 107
 DB 85 SLQKRGIVEOCCTSIICSLYOLENYCN 110

RESULT 12
 Q6YK33 PRELIMINARY; PRT; 110 AA.

AC 06YK33; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Insulin.
 GN Name=INS;
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878;
 RA Stead J.D., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stead J.D.H., Jeffreys A.J.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Belongs to the insulin family.
 DR EMBL; AY137500; AAN06935.1; -
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR InterPro; IPR003234; MolIusc_ins.
 DR Pfam; PF00049; Insulin; 1.
 DR PRINTS; PR00277; INSULINB.
 DR PRODOM; PD015657; MolIusc_ins; 1.
 DR SMART; SM00078; IIGF; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 KM Insulin family.
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 45.5%; Score 267; DB 2; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
 QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREARDLVQGVELGGPGAGSLQPLALEG 84
 QY 86 ----RGIVEOCCTSIICSLYOLENYCN 107
 DB 85 SLQKRGIVEOCCTSIICSLYOLENYCN 110

RESULT 13
 AAP35454 PRELIMINARY; PRT; 110 AA.

AC AAP35454; 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Insulin.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kallme N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
 RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
 RA Phelan M., Farmer A.;
 RT "Cloning of human full-length cDNAs in BD Creator (TM) System Donor
 vector.";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BT006308; AAP35454.1; -
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 45.5%; Score 267; DB 2; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREARDLVQGVELGGPGAGSLQPLALEG 84
 QY 86 ----RGIVEOCCTSIICSLYOLENYCN 107
 DB 85 SLQKRGIVEOCCTSIICSLYOLENYCN 110

RESULT 14

ID AAN06935 PRELIMINARY; PRT; 110 AA.
 AC AAN06935;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE Insulin.
 GN INS.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22833521; PubMed=12952878;
 RA Stead J.D., Hurles M.E., Jeffreys A.J.;
 RT "Global haplotype diversity in the human insulin gene region.";
 RL Genome Res. 13:2101-2111(2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stead J.D.H., Jeffreys A.J.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY137500; AAN06935.1; -
 SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 45.5%; Score 267; DB 2; Length 110;
 Best Local Similarity 60.5%; Pred. No. 3.9e-20;
 Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNQHLGSHLVEALYLVCGERGFFYTPKT----- 85
 DB 25 FVNQHLGSHLVEALYLVCGERGFFYTPKTRREARDLVQGVELGGPGAGSLQPLALEG 84
 QY 86 ----RGIVEOCCTSIICSLYOLENYCN 107
 DB 85 SLQKRGIVEOCCTSIICSLYOLENYCN 110

RESULT 15

```

AAN39451
ID AAN39451 PRELIMINARY; PRT; 110 AA.
AC AAN39451;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Insulin.
GN INS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22833521; PubMed=12952878;
RA Stead J.D., Hurles M.E., Jeffreys A.J.;
RT "Global haplotype diversity in the human insulin gene region.";
RL Genome Res. 13:2101-2111(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Stead J.D.H., Jeffreys A.J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY138590; AAN39451.1; -.
SQ SEQUENCE 110 AA; 11981 MW; C2C3B23B85E520E5 CRC64;

Query Match 45.8%; Score 267; DB 2; Length 110;
Best Local Similarity 60.8%; Pred. No. 3.9e-20;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 56 FVNOHLCGSHVREALYVCGERGFFPTPKT----- 85
DB 25 FVNOHLCGSHVREALYVCGERGFFPTPKTRRRAEDLDVQGVELGGPGAGSLQPLALEG 84
QY 86 ----RGIVEOCTSICTSYQLENYCN 107
DB 85 SLOKRGIVROOCTSICTSYQLENYCN 110

```

Search completed: November 2, 2004, 20:20:34
 Job time : 104.644 secs